

Telegram: @math_books

NUMBER THEORY: CONCEPTS AND PROBLEMS

Telegram: @math_books

Telegram: @math_books

NUMBER THEORY: CONCEPTS AND PROBLEMS

Titu Andreescu

Gabriel Dospinescu

Oleg Mushkarov

Library of Congress Control Number: 2017940046

ISBN-10: 0-9885622-0-0

ISBN-13: 978-0-9885622-0-2

© 2017 XYZ Press, LLC

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (XYZ Press, LLC, 3425 Neiman Rd., Plano, TX 75025, USA) and the authors except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden. The use in this publication of tradenames, trademarks, service marks and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

9 8 7 6 5 4 3 2 1

www.awesomemath.org

Cover design by Iury Ulzutuev

FORWARD

PREDĂ MIHĂILESCU

Exercises are in mathematics like a vitalizer: they strengthen and train the elasticity of the mind, teach a variety of successful methods for approaching specific problems, and enrich the professional culture with interesting questions and results. For a good treatment of a theory, examples and exercises are the art of presenting concrete applications, reflecting the strength and potential of the theoretical results. A strong theory explained only by simple exercise often may reduce the motivation of the reader.

At the other end, there is a wide reserve of problems and exercises of elementary looking nature, but requiring vivid mind and familiarity with a good *bag of tricks*, problems of styles which were much developed by the interest that mathematical competition attracted worldwide in the last 50 years. These problems can only loosely be ordered into applications of individual theories of mathematics, their flavor and interest relaying in the way they combine different areas of knowledge with astute techniques of solving. Often, not always, the problems addressed have some deeper interest of their own and can very well be encountered as intermediate steps in the development of mathematical theories. From this perspective, a good culture of problems can be to a mathematician as helpful, as the familiarity with classical situations in chess matches, to a professional chess player: they develop the aptitude to recognize, formulate and solve individual problems that may play a crucial role in theories and proofs of deeper significance.

The book at hand is a powerful collection of competition problems with number theoretical flavor. They are generally grouped according to common aspects, related to topics like *Divisibility*, *GCD and LCM*, *decomposition of polynomials*, *Congruences and p -adic valuations*, etc. And these aspects can be found in the problems discussed in the respective chapter – beware though to expect much connection to the typical questions one would find in an introductory textbook to number theory, at the chapters with the same name. The problems here are innovative findings and questions, and the connection is more often given by the methods used for the solution, than by the very nature of the problem.

Some problems have a simple combinatorial charm of their own, without requiring much more than good observation – for instance (p. 512, N 25), *Find all $m, n, p \in \mathbf{Q}_{>0}$ such that all of the numbers $m + \frac{1}{np}$, $n + \frac{1}{pm}$, $p + \frac{1}{mn}$ are integers.* Others appear even weird at a first glance, like (p. 656, N 8): *For coprime positive integers p, q , prove that:*

$$\sum_{k=0}^{pq-1} (-1)^{\lfloor k/p \rfloor + \lfloor k/q \rfloor} = \begin{cases} 0 & \text{if } pq \text{ is even} \\ 1 & \text{if } pq \text{ is odd} \end{cases} ;$$

or (N 36, p. 543), requiring to show that infinitely many primes are coprime to the terms of the polynomially recursive sequence given by $a_1 = 1$ and $a_{n+1} = (a_n^2 + 1)^2 - a_n^3$. When one then does the homework, one notices that several useful and non trivial notions about floors are required for solving the problem.

The book also contains some basic propositions, which are in big part classical theorems, but also more specialized results, that can be applied for solving further problems. Thus, beyond the spontaneous charm of some of the exercises, most problems are involved and require a good combination of solid understanding of the theoretical basics, with a good experience in problem solving.

Working through the book one learns a lot. Do you want to know more on how large the difference between the product of k consecutive integers and their LCM can become? A series of results will provide an answer – and you will then certainly find also a set of variations of this theme. For primes p , the Fermat quotient $\phi(2) = \frac{2^{p-1}-1}{p} \pmod p$ has a well known development in terms of harmonic sums. But if you want to know higher terms in its p -adic development, you can find them in the chapter on p -adic values. Together with a series of less known, classical congruences of higher order of Wolfenstone, Morley, Ljunggren et. al., this leads to a series of interesting questions and problems.

Not all problems are atomic training subjects; at the contrary, by a good choice of the problems, the authors may group elementary results that lead to remarkable understanding of some fundamental number theoretical functions, like π, σ, τ, ϕ – the prime distribution function, the number of divisors and their sum, and the Euler totient, respectively. Here also, if you want for instance to

understand how it happens that the fibers of the inverse $\phi^{-1}(X)$ of the Euler totient may become indefinitely large, several exercises lead to the understanding of this phenomenon. It will not surprise that among the authors or solvers of the problems presented, one encounters numerous famous mathematicians, from classical to contemporaneous, ranging from Gauss, Lagrange, Euler and Legendre, through V. Lebesgue, Lucas, but also Hurwitz and, unsurprisingly, Erdős and Schinzel: the borders between research mathematics and advanced problem solving are fluid.

This very short and selective overview of the book should have already suggested that the book can be read with various attitudes and expectations, and there is always much to profit from it. The reader may traverse entire chapters of the book and get familiar with the specifics of the posed problems, but should definitely invest the time for trying to solve at least two or three problems alone, each time when working again with this book. In spite of the well structured construction of the book, one can easily jump to chapters or sections of interest – they are to a large extent self-consistent. And if not, good references help to find the necessary facts which were discussed at previous places of the book.

Altogether – while students eager to acquire experience helping to reach outstanding performance in mathematical competitions will profit most from this book, it is certainly a good companion both for professional mathematicians and for any adult with an active interest in mathematics. Each one of them will find it a leisure to read and work over and over again through the problems of this book.

Preda Mihăilescu

Göttingen, May 2017

Mathematisches Institut der Universität Göttingen

E-mail: preda@uni-math.gwdg.de

Contents

Forward	i
1 Introduction	1
2 Divisibility	3
2.1 Basic properties	3
2.1.1 Divisibility and congruences	3
2.1.2 Divisibility and order relation	10
2.2 Induction and binomial coefficients	22
2.2.1 Proving divisibility by induction	22
2.2.2 Arithmetic of binomial coefficients	26
2.2.3 Derivatives and finite differences	34
2.2.4 The binomial formula	38
2.3 Euclidean division	43
2.3.1 The Euclidean division	43
2.3.2 Combinatorial arguments and complete residue systems	47
2.4 Problems for practice	56
3 GCD and LCM	63
3.1 Bézout's theorem and Gauss' lemma	63
3.1.1 Bézout's theorem and the Euclidean algorithm	63
3.1.2 Relatively prime numbers	68
3.1.3 Inverse modulo n and Gauss' lemma	72
3.2 Applications to diophantine equations and approximations . . .	80
3.2.1 Linear diophantine equations	80

3.2.2	Pythagorean triples	83
3.2.3	The rational root theorem	92
3.2.4	Farey fractions and Pell's equation	96
3.3	Least common multiple	113
3.4	Problems for practice	121
4	The fundamental theorem of arithmetic	129
4.1	Composite numbers	129
4.2	The fundamental theorem of arithmetic	134
4.2.1	The theorem and its first consequences	134
4.2.2	The smallest and largest prime divisor	144
4.2.3	Combinatorial number theory	149
4.3	Infinitude of primes	154
4.3.1	Looking for primes in classical sequences	155
4.3.2	Euclid's argument	160
4.3.3	Euler's and Bonse's inequalities	171
4.4	Arithmetic functions	178
4.4.1	Classical arithmetic functions	178
4.4.2	Multiplicative functions	184
4.4.3	Euler's phi function	194
4.4.4	The Möbius function and its applications	206
4.4.5	Application to squarefree numbers	210
4.5	Problems for practice	216
5	Congruences involving prime numbers	225
5.1	Fermat's little theorem	225
5.1.1	Fermat's little theorem and (pseudo-)primality	225
5.1.2	Some concrete examples	230
5.1.3	Application to primes of the form $4k + 3$ and $3k + 2$	238
5.2	Wilson's theorem	244
5.2.1	Wilson's theorem as criterion of primality	244
5.2.2	Application to sums of two squares	252
5.3	Lagrange's theorem and applications	259
5.3.1	The number of solutions of polynomial congruences	259
5.3.2	The congruence $x^d \equiv 1 \pmod{p}$	266

5.3.3	The Chevalley-Warning theorem	272
5.4	Quadratic residues and quadratic reciprocity	278
5.4.1	Quadratic residues and Legendre's symbol	278
5.4.2	Points on spheres mod p and Gauss sums	286
5.4.3	The quadratic reciprocity law	297
5.5	Congruences involving rational numbers and binomial coefficients	304
5.5.1	Binomial coefficients modulo primes: Lucas' theorem . .	304
5.5.2	Congruences involving rational numbers	310
5.5.3	Higher congruences: Fleck, Morley, Wolstenholme,... .	316
5.5.4	Hensel's lemma	324
5.6	Problems for practice	330
6	p-adic valuations and the distribution of primes	341
6.1	The yoga of p -adic valuations	341
6.1.1	The local-global principle	341
6.1.2	The strong triangle inequality	347
6.1.3	Lifting the exponent lemma	353
6.2	Legendre's formula	360
6.2.1	The p -adic valuation of $n!$: the exact formula	360
6.2.2	The p -adic valuation of $n!$: inequalities	363
6.2.3	Kummer's theorem	369
6.3	Estimates for binomial coefficients and the distribution of prime numbers	373
6.3.1	Central binomial coefficients and Erdős' inequality . . .	373
6.3.2	Estimating $\pi(n)$	376
6.3.3	Bertrand's postulate	380
6.4	Problems for practice	386
7	Congruences for composite moduli	393
7.1	The Chinese remainder theorem	393
7.1.1	Proof of the theorem and first examples	393
7.1.2	The local-global principle	400
7.1.3	Covering systems of congruences	408
7.2	Euler's theorem	417

7.2.1	Reduced residue systems and Euler's theorem	417
7.2.2	Practicing Euler's theorem	421
7.3	Order modulo n	427
7.3.1	Elementary properties and examples	427
7.3.2	Practicing the notion of order modulo n	440
7.3.3	Primitive roots modulo n	448
7.4	Problems for practice	460
8	Solutions to practice problems	467
8.1	Divisibility	467
8.2	GCD and LCM	496
8.3	The fundamental theorem of arithmetic	523
8.4	Congruences involving prime numbers	568
8.5	p -adic valuations and the distribution of primes	620
8.6	Congruences for composite moduli	652
	Bibliography	683
	Other Books from XYZ Press	685

Chapter 1

Introduction

Based on lectures given by the authors at the AwesomeMath Summer Program over several years, this book is a slightly non-standard introduction to elementary number theory. Nevertheless, it still develops theoretical concepts from scratch with full proofs. The book insists on exemplifying these results through interesting and rather challenging problems. In particular, the reader will not find many advanced concepts in this book, but will encounter quite a lot of intriguing results that can be proven using “basic” number theory yet nonetheless test one’s problem-solving aptitude.

The book is divided into six large chapters, each focusing on a fundamental concept or result. Each chapter is itself divided into sections that reinforce a specific topic through a large series of examples arranged (subjectively) in increasing order of difficulty. In particular, the first two chapters are largely elementary but fundamental for appreciating the rest of the book. The topics explored in these two chapters are classical: divisibility, congruences, Euclidean division, greatest common divisor, and least common multiple. With the theoretical concepts being fairly elementary, the focus is more on concrete problems and interesting applications, for instance, Diophantine equations, finite differences, and problems with a combinatorial flavor. The third chapter is devoted to the fundamental theorem of arithmetic and its numerous applications. After proving basic properties of prime numbers and the uniqueness of prime factorization, the authors emphasize their utility and vast scope among

arithmetic functions. There are many non-standard and sometimes surprising results in this chapter.

The fourth and fifth chapters, devoted to congruences involving prime numbers and to the distribution of prime numbers, are in some sense the heart of the book. Each of the classical congruences (Fermat, Wilson, Lagrange, and Lucas) is studied in depth in the fourth chapter, along with numerous examples of their use, for instance, quadratic residues, the number of solutions to polynomial congruences, and congruences involving binomial coefficients or higher congruences. In the fifth chapter, p -adic valuations are used to study the distribution of prime numbers. This has the advantage of being fairly elementary, while still producing beautiful and nontrivial results. The key results of this chapter are Legendre's theorem and the arithmetic of binomial coefficients, leading to strong results concerning the distribution of prime numbers. Finally, the sixth chapter discusses congruences for composite moduli, introducing further essential concepts and results: the Chinese remainder theorem, Euler's theorem, and their applications to primitive roots modulo integers. The main focus is again providing many examples of these concepts' applications (in particular, the reader will find a whole section devoted to systems of congruences). Each chapter contains a long list of practice problems, whose solutions are presented at the end of the book.

Experience has shown that it is easier to make students appreciate the beauty and power of a result when it is enhanced by pertinent and challenging examples. We strove to achieve this, a possible explanation for the book's length, although the theoretical material is rather classical and standard.

We would like to thank our students at the AwesomeMath Summer Program on whom we tested a large part of this material and who supplied many of the solutions presented here. We are also indebted to Richard Stong for a very careful reading of the book, for pointing out many inaccuracies, and for supplying a great deal of solutions (many of which were simpler and more elegant than ours!).

Chapter 2

Divisibility

This first chapter is fairly elementary and discusses basic properties of divisibility, congruences and the Euclidean division. These will be constantly used later on in the book and represent the foundations of arithmetic, on which we will build more advanced results later on. We tried to insist more on relatively nonstandard examples or applications, some of which are relatively nontrivial (such as the topic of finite differences and their applications to congruences).

2.1 Basic properties

In this section we introduce the notion of divisibility and study some of its basic properties.

2.1.1 Divisibility and congruences

We start by defining the divisibility relation.

Definition 2.1. Let a, b be integers. We say that a divides b and write $a \mid b$ if there is an integer c such that $b = ac$.

There are many equivalent ways of saying that a divides b : we can also say that b is divisible by a , that a is a divisor of b or that b is a multiple of a . All

these formulations are used in practice. Note that if $a \neq 0$, then saying that a divides b is equivalent to saying that the rational number $\frac{b}{a}$ is an integer. The previous definition takes into account the possibility that $a = 0$, in which case a divides b if and only if $b = 0$. In other words, any integer is a divisor of 0, and 0 is the only multiple of 0.

If 2 divides an integer n , we say that n is even. Otherwise, we say that n is odd. Thus the even integers are $\dots, -2, 0, 2, 4, 6, \dots$, while the odd ones are $\dots -3, -1, 1, 3, 5, \dots$. Note that if n is odd, then $n - 1$ is even, in other words any integer n is either of the form $2k$ or $2k + 1$ for some integer k . In particular, we obtain that the product of two consecutive integers is always even. We deduce for instance that if a is an odd integer, say $a = 2k + 1$, then

$$a^2 - 1 = 4k(k + 1)$$

is a multiple of 8. In particular any perfect square (i.e. number of the form x^2 with x an integer) is either a multiple of 4 or of the form $8k + 1$ for some integer k .

The following result summarizes the basic properties of the divisibility relation.

Proposition 2.2. *The divisibility relation has the following properties:*

1. (reflexivity) a divides a for all integers a .
2. (transitivity) If $a \mid b$ and $b \mid c$, then $a \mid c$.
3. If a, b_1, \dots, b_n are integers and $a \mid b_i$ for $1 \leq i \leq n$, then $a \mid b_1c_1 + \dots + b_nc_n$ for all integers c_1, \dots, c_n .
4. If $a \mid b$ and $a \mid b \pm c$, then $a \mid c$.
5. If $n \mid a - b$ and $n \mid a' - b'$, then $n \mid aa' - bb'$.

Proof. All of these properties follow straight from the definition. We only prove properties 3) and 5) here, leaving the others to the reader. For property 3), write $b_i = ax_i$ for some integers x_i . Then

$$b_1c_1 + \dots + b_nc_n = ax_1c_1 + \dots + ax_nc_n = a(x_1c_1 + \dots + x_nc_n)$$

is a multiple of a . For property 5), write $a - b = kn$ and $a' - b' = k'n$ for some integers k, k' . Then

$$aa' - bb' = (b + kn)(b' + k'n) - bb' = n(bk' + b'k + nkk'),$$

thus $n \mid aa' - bb'$. □

We introduce next a key notation and definition, that of congruences:

Definition 2.3. Let a, b, n be integers. We say that a and b are congruent modulo n and write

$$a \equiv b \pmod{n}$$

if $n \mid a - b$.

Most parts of the following theorem are simple reinterpretations of proposition 2.2. They are of constant use in practice.

Theorem 2.4. For all integers a, b, c, d, n we have

- a) (reflexivity) $a \equiv a \pmod{n}$.
- b) (symmetry) If $a \equiv b \pmod{n}$ then $b \equiv a \pmod{n}$.
- c) (transitivity) If $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$, then $a \equiv c \pmod{n}$.
- d) If $a \equiv c \pmod{n}$ and $b \equiv d \pmod{n}$, then $a + b \equiv c + d \pmod{n}$ and $ab \equiv cd \pmod{n}$.
- e) If $a \equiv b \pmod{n}$, then $ac \equiv bc \pmod{nc}$. Conversely, if $ac \equiv bc \pmod{nc}$ and $c \neq 0$, then $a \equiv b \pmod{n}$.

Proof. a), b), c), d) are either clear or consequences of proposition 2.2. Property e) is immediate and left to the reader. □

Remark 2.5. We cannot cancel congruences without taking care. In other words, it is not true that if $ab \equiv ac \pmod{n}$, then $b \equiv c \pmod{n}$ or $a \equiv 0 \pmod{n}$. For instance $2 \cdot 2 \equiv 2 \cdot 0 \pmod{4}$, but 2 is not congruent to 0 modulo 4. We will see later on that we can "cancel a " in a congruence $ab \equiv ac \pmod{n}$ provided n and a share no common divisor except ± 1 .

Let us illustrate the previous theorem with some concrete problems (where no congruence is mentioned!).

Example 2.6. Find the last digit of $9^{1003} - 7^{902} + 3^{801}$.

Proof. We have $9^{1003} \equiv (-1)^{1003} \equiv -1 \equiv 9 \pmod{10}$. In addition,

$$7^{902} \equiv 49^{451} \equiv (-1)^{451} \equiv -1 \pmod{10}.$$

Finally,

$$3^{801} \equiv 3 \cdot (3^4)^{200} \equiv 3 \cdot 1^{200} \equiv 3 \pmod{10}.$$

Hence

$$9^{1003} - 7^{902} + 3^{801} \equiv (-1) - (-1) + 3 \equiv 3 \pmod{10},$$

so the last digit is 3. □

Example 2.7. Prove that for any $n \in \mathbf{N}$ the number $a_n = 11^{n+2} + 12^{2n+1}$ is divisible by 133.

Proof. We have $12^2 = 144 \equiv 11 \pmod{133}$, hence

$$a_n \equiv 11^{n+2} + 12 \cdot 144^n \equiv 11^{n+2} + 12 \cdot 11^n \equiv 11^n(121 + 12) \equiv 0 \pmod{133}. \quad \square$$

Example 2.8. (Kvant, M 274) Find the least number of the form:

- (i) $|11^k - 5^l|$,
- (ii) $|36^k - 5^l|$,
- (iii) $|53^k - 37^l|$,

where k and l are positive integers.

Proof. (i) The last digit of $|11^k - 5^l|$ is either 6 or 4, thus the least number of the form $|11^k - 5^l|$ must be at least 4. Since $|11^2 - 5^3| = 4$, we deduce that the answer is 4.

(ii) We have $11 = |36 - 5^2|$ and we will show that this is the least number of the form $|36^k - 5^l|$. Suppose that for some k, l we have $|36^k - 5^l| \leq 10$. Since $36^k - 5^l \equiv 6 - 5 = 1 \pmod{10}$, we deduce that $36^k - 5^l = 1$ or $36^k - 5^l = -9$. The first equality is impossible since it would imply that $0 - 1 \equiv 1 \pmod{4}$, impossible. The second equality is also impossible since it would yield $0 - (-1)^l \equiv 0 \pmod{3}$, again impossible. This finishes the proof.

(iii) Note first that the given numbers are divisible by 4 since 53^k and 37^l are congruent to 1 modulo 4. We will show that the desired number is $16 = |53 - 37|$. Note that

$$53^k \equiv (-1)^k \pmod{9}, \quad 37^l \equiv 1 \pmod{9}.$$

Hence $N = |53^k - 37^l| \equiv 0, \pm 2 \pmod{9}$ which shows that $N \neq 4, 8, 12$. \square

The following fundamental theorem is of constant use.

Theorem 2.9. a) If a, b are integers, then $a - b \mid a^k - b^k$ for all $k \geq 1$.

b) More generally, if d, n are positive integers such that $d \mid n$, then $a^d - b^d \mid a^n - b^n$ for all integers a, b . Moreover, if $\frac{n}{d}$ is odd, then $a^d + b^d \mid a^n + b^n$ for all integers a, b (in particular $a + b \mid a^n + b^n$ for all integers a, b if n is odd).

Proof. a) This follows directly from the identity

$$a^k - b^k = (a - b)(a^{k-1} + a^{k-2}b + \dots + ab^{k-2} + b^{k-1}).$$

b) Let $n = kd$ for some positive integer k . Then setting $x = a^d$, $y = b^d$ we are reduced to showing that $x - y \mid x^k - y^k$ (which follows from part a)) and $x + y \mid x^k + y^k$ when k is odd, which follows from

$$x + y = x - (-y) \mid x^k - (-y)^k = x^k + y^k. \quad \square$$

Remark 2.10. 1) We will see later on that under rather weak hypotheses, the divisibility $a^m - b^m \mid a^n - b^n$ implies $m \mid n$.

2) The identity

$$a^n - b^n = (a - b)(a^{n-1} + a^{n-2}b + \dots + ab^{n-2} + b^{n-1})$$

is absolutely fundamental in arithmetic and the reader should become very familiar with it, since it will be used constantly in this book. Indeed, in many cases the results of theorem 2.9 are strong enough, but in some circumstances a finer analysis of the term $a^{n-1} + a^{n-2}b + \dots + b^{n-1}$ is crucial.

The following result is a simple translation of the previous theorem in terms of congruences:

Corollary 2.11. *Let a, b, n be integers, let k be a positive integer and let $d \mid k$ a positive divisor of k .*

- a) *If $a \equiv b \pmod{n}$, then $a^k \equiv b^k \pmod{n}$.*
- b) *If $a^d \equiv b^d \pmod{n}$, then $a^k \equiv b^k \pmod{n}$.*
- c) *If $a^d \equiv -b^d \pmod{n}$ and $\frac{k}{d}$ is odd, then $a^k \equiv -b^k \pmod{n}$.*

Example 2.12. Using that $641 = 2^7 \cdot 5 + 1$, prove that $641 \mid 2^{32} + 1$.

Proof. We have $2^7 \cdot 5 \equiv -1 \pmod{641}$, thus $2^{28} \cdot 5^4 \equiv 1 \pmod{641}$. Since $641 = 5^4 + 2^4$ we have $5^4 \equiv -2^4 \pmod{641}$, thus $2^{28} \cdot 5^4 \equiv -2^{32} \pmod{641}$ and so $-2^{32} \equiv 1 \pmod{641}$, which is exactly what we need. \square

Example 2.13. a) Prove that if n is a positive integer, then 9 divides the difference between n and the sum of its decimal digits.

b) Let n be a positive integer and let S_1 (respectively S_2) be the sum of the digits of n at the odd (respectively even) positions (the last digit of n has position 0). Prove that $n \equiv S_2 - S_1 \pmod{11}$.

Proof. a) Write

$$n = \overline{a_k a_{k-1} \dots a_0} = a_k \cdot 10^k + a_{k-1} 10^{k-1} + \dots + a_0$$

for some decimal digits a_k, \dots, a_0 with $a_k \neq 0$. Then

$$n - (a_0 + a_1 + \dots + a_k) = a_k(10^k - 1) + a_{k-1}(10^{k-1} - 1) + \dots + a_1(10 - 1)$$

is a multiple of 9, since each term in the sum is a multiple of 9 thanks to theorem 2.9.

b) The proof is identical to that of part a), the key point being the congruence $10^i \equiv (-1)^i \pmod{11}$ for all i . \square

Example 2.14. (Kvant M 676) Prove that for every positive integer n the sum of the digits of 1981^n is not less than 19.

Proof. Write $S(x)$ for the sum of the decimal digits of x . Since $9 \mid x - S(x)$ for all x and since $9 \mid 1981^n - 1$ (as $9 \mid 1980$), it follows that $9 \mid S(1981^n) - 1$ and so $S(1981^n)$ is one of the numbers 1, 10, 19, Since 1981^n ends in 1 (because

$10 \mid 1981^n - 1$) it follows that $S(1981^n) > 1$. Suppose that $S(1981^n) = 10$, thus $S(1981^n - 1) = 9$. Denote by S_1 (respectively S_2) the sum of the digits of $1981^n - 1$ at the odd (respectively even) positions. Then $0 \leq S_1, S_2 \leq 9$. On the other hand $1981^n - 1$ is divisible by 1980, thus it is divisible by 11. Hence $S_1 - S_2$ is divisible by 11 (by the previous example) and we conclude that $S_1 = S_2$. But $S_1 + S_2 = 9$, a contradiction. Thus $S(1981^n) \geq 19$ for all n . \square

Example 2.15. Let $F_n = 2^{2^n} + 1$ be the n th Fermat number. Prove that $F_n \mid 2^{F_n} - 2$ for all $n \geq 1$.

Proof. It suffices to show that $F_n \mid 2^{F_n-1} - 1$. Note that

$$F_n \mid (2^{2^n} - 1)(2^{2^n} + 1) = 2^{2^{n+1}} - 1.$$

If $a \mid b$ then $2^a - 1 \mid 2^b - 1$ by theorem 2.9. It suffices therefore to show that $2^{n+1} \mid F_n - 1$, or equivalently $n + 1 \leq 2^n$. This is clear. \square

An immediate consequence of the previous theorem is the following very useful:

Proposition 2.16. *If f is a polynomial with integer coefficients, then for all integers a, b*

$$a - b \mid f(a) - f(b).$$

Thus, if $a \equiv b \pmod{n}$ for some integer n , then $f(a) \equiv f(b) \pmod{n}$.

Proof. Write

$$f(X) = c_0 + c_1X + \dots + c_nX^n$$

for some integers c_0, \dots, c_n and some $n \geq 0$. Then

$$f(a) - f(b) = c_1(a - b) + c_2(a^2 - b^2) + \dots + c_n(a^n - b^n)$$

and each term in the sum is a multiple of n by theorem 2.9. The result follows. \square

Example 2.17. Let f be a polynomial with integer coefficients and let a be a positive integer such that $f(a) \neq 0$. Prove that there are infinitely many positive integers b such that $f(a) \mid f(b)$.

Proof. We take $b = a + k|f(a)|$ with k a positive integer. Then

$$f(a) \mid k|f(a)| = b - a \mid f(b) - f(a)$$

and so $f(a) \mid f(b)$. Since k is arbitrary, the result follows. \square

2.1.2 Divisibility and order relation

Another key property of the divisibility relation that we want to emphasize in this section is its relationship with the usual order on the set of integers: the next proposition roughly says that a divisor of a number cannot exceed that number. One has to be a little bit careful when making such a statement (note that 1 is a divisor of -2 , but it is certainly not less than -2), so we formalize this as follows:

Proposition 2.18. *If a divides b and $b \neq 0$, then $|a| \leq |b|$.*

Proof. Write $b = ac$, then $c \neq 0$ (since $b \neq 0$), hence $|b| = |a| \cdot |c| \geq |a|$. \square

Remark 2.19. The hypothesis $b \neq 0$ is crucial in the previous proposition. The number 0 plays a very special role: it is the only integer having infinitely many divisors. More precisely, 0 is divisible by all integers, since if a is any integer, then $0 = a \cdot 0$. On the other hand, if $n \in \mathbf{Z}$ has infinitely many divisors, then necessarily $n = 0$: otherwise, by the previous proposition any divisor d of n satisfies $d \in \{-|n|, \dots, 0, 1, \dots, |n|\}$, hence n has only finitely many divisors. The next example is a nice illustration of this important observation.

Example 2.20. (Russia 1964) Let a, b be integers and let n be a positive integer such that $k - b \mid k^n - a$ for infinitely many integers k . Prove that $a = b^n$.

Proof. For any integer k we have $k - b \mid k^n - b^n$, so if $k - b \mid k^n - a$, then

$$k - b \mid (k^n - b^n) - (k^n - a) = a - b^n.$$

Using the hypothesis of the problem, we deduce that $a - b^n$ has infinitely many divisors and so $a - b^n = 0$. The result follows. \square

One of the consequences of the previous proposition is the following property of the divisibility relation.

Corollary 2.21. *If a, b are integers such that $a \mid b$ and $b \mid a$, then $|a| = |b|$, i.e. $a = \pm b$.*

Proof. Everything is clear if $a = 0$ or $b = 0$. Otherwise, the previous proposition gives $|a| \leq |b|$ and $|b| \leq |a|$, thus $|a| = |b|$. \square

Example 2.22. Find all integers n such that $a - b \mid a^2 + b^2 - nab$ for all distinct integers a, b .

Proof. The identity $a^2 + b^2 - nab = (a - b)^2 + (2 - n)ab$ shows that $a - b \mid (2 - n)ab$ for all $a \neq b \in \mathbb{Z}$. Taking $b = 1$ and $a = k + 1$, with k a positive integer, we deduce that $k \mid (2 - n)(k + 1) = (2 - n)k + 2 - n$ and so $k \mid 2 - n$. Hence $2 - n$ has infinitely many divisors and $n = 2$. Conversely, $n = 2$ is a solution of the problem. \square

Example 2.23. (Putnam 2007) Let f be a nonconstant polynomial with positive integer coefficients. Prove that if n is a positive integer, then $f(n)$ divides $f(f(n) + 1)$ if and only if $n = 1$.

Proof. We have $f(f(n) + 1) \equiv f(1) \pmod{f(n)}$. If $n = 1$, then this implies that $f(f(n) + 1)$ is divisible by $f(n)$. Otherwise, $0 < f(1) < f(n)$ since f is nonconstant and has positive coefficients, so $f(f(n) + 1)$ cannot be divisible by $f(n)$. \square

Example 2.24. a) Prove that for any positive integer n there are distinct positive integers x and y such that $x + j$ divides $y + j$ for $j = 1, 2, 3, \dots, n$.

b) Suppose that x, y are positive integers such that $x + j$ divides $y + j$ for all positive integers j . Prove that $x = y$.

Proof. a) We have $x + j \mid y + j$ if and only if $x + j \mid (y + j) - (x + j) = y - x$. Thus it is enough to ensure that $y - x$ is a multiple of $(x + 1)(x + 2) \dots (x + n)$, for instance $y = x + (x + 1)(x + 2) \dots (x + n)$.

b) Arguing as in a), we see that $y - x$ must be a multiple of $x + j$ for all positive integers j . Remark 2.19 yields $y = x$ and we are done. \square

Example 2.25. Let f be a polynomial with integer coefficients, of degree $n > 1$. What is the maximal number of consecutive integers belonging to the sequence $f(1), f(2), f(3), \dots$?

Proof. For the polynomial $f(X) = X + (X - 1)(X - 2)\dots(X - n)$ we have $f(1) = 1, f(2) = 2, \dots, f(n) = n$, thus we have n consecutive numbers in the sequence $f(1), f(2), \dots$. We will prove that we cannot have more. Assume for contradiction that we can find positive integers a_1, \dots, a_{n+1} and an integer x such that $f(a_i) = x + i$ for $1 \leq i \leq n + 1$. Then $f(a_{i+1}) - f(a_i) = 1$ is a multiple of $a_{i+1} - a_i$, thus $a_{i+1} - a_i$ equals 1 or -1 for all i . Since a_1, \dots, a_{n+1} are clearly pairwise distinct (since so are their images by f), we deduce that we cannot have sign changes in the sequence $a_2 - a_1, a_3 - a_2, \dots, a_{n+1} - a_n$ (indeed, otherwise there would exist i such that $a_{i+1} - a_i$ is the opposite of $a_{i+2} - a_{i+1}$, which would force $a_i = a_{i+2}$). Thus the sequence $a_2 - a_1, a_3 - a_2, \dots, a_{n+1} - a_n$ must either consist only of 1's or only of -1 's. We can thus find a sign ε such that $a_{i+1} - a_i = \varepsilon$ for all i . But then $a_i = a_1 + \varepsilon \cdot (i - 1)$ for all i , hence $f(a_1 - \varepsilon + \varepsilon \cdot i) = x + i$ for $1 \leq i \leq n + 1$. We deduce that the polynomial $f(a_1 - \varepsilon + \varepsilon \cdot X) - x - X$ has at least $n + 1$ distinct roots, which is impossible since it has degree precisely n . This proves that the answer of the problem is n . \square

Example 2.26. Let f be a polynomial with integer coefficients, of degree $n \geq 2$. Prove that the equation $f(f(x)) = x$ has at most n integral solutions.

Proof. Let x, y be distinct integers such that $f(f(x)) = x$ and $f(f(y)) = y$. Then $x - y = f(f(x)) - f(f(y))$ is a multiple of $f(x) - f(y)$, which in turn is a multiple of $x - y$. Thus necessarily $|f(x) - f(y)| = |x - y|$. Consider now integers $a_1 < \dots < a_d$ such that $f(f(a_i)) = a_i$ for $1 \leq i \leq d$. Then the previous observation yields $|f(a_i) - f(a_j)| = a_j - a_i$ for $i < j$. We claim that the sequence $f(a_1), \dots, f(a_d)$ is either increasing or decreasing. Indeed, we have

$$\begin{aligned} |f(a_{i+1}) - f(a_i) + f(a_{i+2}) - f(a_{i+1})| &= |f(a_{i+2}) - f(a_i)| \\ &= a_{i+2} - a_i = |f(a_{i+1}) - f(a_i)| + |f(a_{i+2}) - f(a_{i+1})|, \end{aligned}$$

therefore $f(a_{i+1}) - f(a_i)$ and $f(a_{i+2}) - f(a_{i+1})$ must have the same sign for all i , proving the claim.

Assume that $f(a_1), \dots, f(a_n)$ is increasing (the other case is similar). Then necessarily $f(a_{i+1}) - f(a_i) = a_{i+1} - a_i$ for all i , in other words there is some

number c such that $f(a_i) - a_i = c$ for $1 \leq i \leq d$. Since $f(X) - X - c$ has degree n , it can have at most n distinct roots and so $d \leq n$, as desired. \square

Remark 2.27. A more general problem (in which $f \circ f$ is replaced with $f \circ f \circ \dots \circ f$) was proposed at the IMO 2006.

Example 2.28. (Tournament of the Towns 2002) Let $a_1 < a_2 < \dots$ be an infinite increasing sequence of positive integers such that a_n divides $a_1 + a_2 + \dots + a_{n-1}$ for $n \geq 2002$. Prove that there is a positive integer n_0 such that

$$a_n = a_1 + \dots + a_{n-1}$$

for all $n \geq n_0$.

Proof. By hypothesis, there is a sequence $x_{2002}, x_{2003}, \dots$ of positive integers such that for all $n \geq 2002$ we have

$$a_1 + a_2 + \dots + a_{n-1} = x_n a_n.$$

Write the previous relation with $n + 1$ instead of n and subtract the two resulting relations. We obtain

$$x_{n+1} a_{n+1} = x_n a_n + a_n = a_n (x_n + 1) \quad (1)$$

We deduce that

$$x_{n+1} = \frac{a_n}{a_{n+1}} (x_n + 1) < x_n + 1,$$

since $a_n < a_{n+1}$. Consequently, $x_{n+1} \leq x_n$ for $n \geq 2002$. Since there is no decreasing infinite sequence of positive integers, we deduce that there is $n_0 \geq 2002$ such that for all $n \geq n_0$ we have $x_{n+1} = x_n$. Let $k = x_{n_0}$, then $x_n = k$ for $n \geq n_0$ and relation (1) becomes

$$k a_{n+1} = (k + 1) a_n$$

for $n \geq n_0$. In particular,

$$a_n = k(a_{n+1} - a_n)$$

is a multiple of k for $n \geq n_0$. Writing $a_n = kb_n$, we also have $b_n = k(b_{n+1} - b_n)$ and so $k \mid b_n$ for all n , that is $k^2 \mid a_n$ for all $n \geq n_0$. An immediate induction then shows that $k^j \mid a_n$ for all $j \geq 1$ and all $n \geq n_0$. In particular, $k^j \leq a_{n_0}$ for all $j \geq 1$, which forces $k = 1$. But then

$$a_1 + \dots + a_{n-1} = ka_n = a_n$$

for $n \geq n_0$ and we are done. \square

A fundamental property that easily follows from the relationship between divisibility and order relation as well as basic properties of odd and even numbers is:

Theorem 2.29. *Let n be a nonzero integer. There is a unique pair of integers (a, b) with $a \geq 0$, b odd and $n = 2^a \cdot b$.*

Proof. Let us start by proving uniqueness. Suppose that $2^a b = 2^c d$ with $a, c \geq 0$ and b, d odd, and assume that $a \neq c$. Without loss of generality, we may assume that $a < c$, then $b = 2^{c-a} d$ is even, a contradiction. Thus $a = c$ and then $b = d$.

In order to prove the existence part, consider the set of powers of 2 which divide n . This set is finite, since if 2^a divides n , then $a < 2^a \leq |n|$. Thus there is a largest integer a such that $2^a \mid n$. Write $n = 2^a b$ for some integer b . If b is even, then $b = 2c$ for some integer c and then $2^{a+1} \mid n$, contradicting the maximality of a . Thus b is odd and the result is proved. \square

Remark 2.30. 1) It follows easily from the previous theorem that if a, b are integers such that ab is a power of 2, i.e. $ab = 2^n$ for some $n \geq 0$ then $|a|$ and $|b|$ (but not necessarily a and b) are also powers of 2.

2) From the uniqueness part of the theorem, it follows that if $n = 2m$ is even and an odd number d divides n , then d divides m . This is our first example of a cancellation in congruences and we will use it frequently.

Yet another result that is fairly useful in practice is the following:

Theorem 2.31. *If a is an odd integer, then for all $n \geq 0$*

$$2^{n+2} \mid a^{2^n} - 1.$$

Proof. We have

$$a^{2^n} - 1 = (a - 1)(a + 1)(a^2 + 1)(a^4 + 1) \dots (a^{2^{n-1}} + 1).$$

Since a is odd, $(a - 1)(a + 1) = a^2 - 1$ is a multiple of 8, and $a^2 + 1, a^4 + 1, \dots, a^{2^{n-1}} + 1$ are each multiples of 2. Hence $a^{2^n} - 1$ is a multiple of $2^{3+(n-1)} = 2^{n+2}$, as desired.

Of course, the statement can also be proved by induction on n : for $n = 0$ it is equivalent to $8 \mid a^2 - 1$, which we have already seen. Assuming that $a^{2^n} = 1 + k \cdot 2^{n+2}$, we have

$$a^{2^{n+1}} = (a^{2^n})^2 = (1 + k \cdot 2^{n+2})^2 = 1 + k \cdot 2^{n+3} + k^2 2^{2n+4} = 1 + (k + k^2 2^{n+1}) 2^{n+3}$$

and the result follows. \square

We will use the previous two theorems throughout the book. The next examples are a few illustrations of these results.

Example 2.32. Let n be an integer greater than 1. Prove that n is odd if and only if n divides $1^n + 2^n + \dots + (n - 1)^n$.

Proof. If n is odd, then $k^n + (n - k)^n$ is a multiple of n for $1 \leq k \leq n - 1$, hence $2(1^n + 2^n + \dots + (n - 1)^n)$ is a multiple of n and then $n \mid 1^n + 2^n + \dots + (n - 1)^n$. Suppose that n is even and write $n = 2^a m$ with $a \geq 1$ and m odd. If k is odd, then $k^n = (k^{2^a})^m \equiv 1 \pmod{2^a}$, while if k is even, then $k^n \equiv 0 \pmod{2^a}$. We deduce that

$$1^n + 2^n + \dots + (n - 1)^n \equiv 2^{a-1} m \pmod{2^a}$$

and so 2^a cannot divide $1^n + 2^n + \dots + (n - 1)^n$. \square

Example 2.33. Prove that if $n > 1$ then $s = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n}$ is not an integer.

Proof. Let a be the product of all odd integers less than or equal to n , and let k be the largest integer such that $2^k \leq n$. We claim that $2^{k-1}as$ is not an integer. If $1 \leq m \leq n$ with $m \neq 2^k$, then m can be written in the form $2^t u$ where $0 \leq t \leq k - 1$ and $1 \leq u \leq n$ is an odd integer. Hence $m \mid 2^{k-1}a$, so $2^{k-1}a \cdot \frac{1}{m}$ is an integer. Hence $2^{k-1}as = N + \frac{a}{2}$ for some integer N . But a is odd, hence $\frac{a}{2}$ is not an integer. It follows that s is not an integer. \square

Example 2.34. Is there a polynomial $f(x, y)$ in two variables, with integer coefficients and having the following properties:

- a) The equation $f(x, y) = 0$ has no integral solutions.
- b) For each positive integer n there are integers x, y such that $n \mid f(x, y)$?

Proof. We will show that $f(x, y) = (2x - 1)(3y - 1)$ is such a polynomial. It is clear that $f(x, y) = 0$ has no integral solutions, since it forces $x = \frac{1}{2}$ or $y = \frac{1}{3}$. Now let n be a positive integer and write $n = 2^k m$ with $k \geq 0$ and m odd. Note that $3 \mid 2^{2k+1} + 1 = 2 \cdot 4^k + 1$ (since $3 \mid 4^k - 1$), thus we can write $2^{2k+1} = 3y - 1$ for some integer y . Setting $x = \frac{m+1}{2}$ (an integer, since m is odd) we obtain

$$(2x - 1)(3y - 1) = m \cdot 2^{2k+1},$$

a multiple of n . □

Example 2.35. (Turkey TST 2016) Find all functions $f : \mathbb{N} \rightarrow \mathbb{N}$ such that for all $m, n \geq 1$ we have $f(mn) = f(m)f(n)$ and $m + n \mid f(m) + f(n)$.

Proof. Clearly for any odd positive integer k setting $f(x) = x^k$ yields a solution of the problem. We will prove that these are all solutions. First of all, note that $f(1) = 1$ since $f(1) = f(1)^2$ and $f(1)$ is positive. Next, we focus on $f(2)$. Write $f(2) = 2^k(2r + 1)$ for some $k, r \geq 0$. Assume that $r > 0$, then

$$1 + 2r \mid f(1) + f(2r) = 1 + f(2)f(r) = 1 + 2^k(2r + 1)f(r),$$

thus $1 + 2r \mid 1$, impossible. Thus $f(2) = 2^k$. Since $f(mn) = f(m)f(n)$ for all m, n , we have $f(2^n) = 2^{nk}$ for all $n \geq 1$. Since $6 \mid f(2) + f(4) = 2^k + 4^k$, we deduce that k is odd. Finally, for any $n \geq 1$ and any $d \geq 1$ we have $n + 2^d \mid f(n) + f(2^d) = f(n) + 2^{kd}$ and, since k is odd, we also have $n + 2^d \mid n^k + 2^{kd}$, thus $n + 2^d \mid f(n) - n^k$. Fixing n and letting d vary we deduce that $f(n) = n^k$ for all n (since $f(n) - n^k$ has infinitely many divisors, namely all numbers $n + 2^d$ with $d \geq 1$), finishing the proof. □

We end this section with a few more challenging examples, which combine most of the time all the previous techniques. The first one is a famous IMO problem, which became an absolute classic. The method of the proof (also known as infinite descent) goes back to Fermat and crucially uses the ordering

of integers. In many cases we need to prove that certain diophantine equations $f(x_1, \dots, x_k) = 0$ have no solutions, or only "trivial solutions" (those that one can find "at a glance"). The idea is to start with a potential solution (or a "nontrivial" solution) of the equation and produce a "smaller" one. If the "smaller" solution created is not "trivial", one repeats the process. We obtain this way a sequence of solutions, which become "smaller" and "smaller", forcing therefore the process to stop. One can also argue (perhaps more directly) by contradiction and consider a "minimal" solution of the problem and then reach a contradiction after having created a "smaller" solution.

Let's see how this works precisely in a simple example, before embarking in the more challenging example below. Consider the equation $x^2 + y^2 = 3z^2$. We claim that the only solution in integers is the "trivial" one, namely $x = y = z = 0$. Indeed, consider a solution (x, y, z) which is not trivial. Then 3 divides $x^2 + y^2$. Checking several cases, it is a simple matter to deduce that x, y must both be multiples of 3. Then $x = 3x_1, y = 3y_1$ and $z^2 = 3(x_1^2 + y_1^2)$. We deduce that z is a multiple of 3 (otherwise $z^2 \equiv 1 \pmod{3}$), say $z = 3z_1$ and then $x_1^2 + y_1^2 = 3z_1^2$. Thus (x_1, y_1, z_1) is also a solution of the equation, and it is not trivial, since (x, y, z) is not trivial. On the other hand

$$|x_1| + |y_1| + |z_1| = \frac{|x| + |y| + |z|}{3} < |x| + |y| + |z|,$$

thus the solution (x_1, y_1, z_1) is "smaller" than (x, y, z) , in the sense that the sum of the absolute values of x_1, y_1, z_1 is smaller than that of x, y, z . Considering a nontrivial solution (x, y, z) with $|x| + |y| + |z|$ minimal, this immediately yields a contradiction.

Example 2.36. (IMO 1988) Let a, b be positive integers such that $ab+1$ divides $a^2 + b^2$. Prove that $\frac{a^2+b^2}{ab+1}$ is a perfect square.

Proof. Assume that this fails for some a, b and pick a pair (a, b) for which this fails and for which $a + b$ has the smallest possible value. Write $a^2 + b^2 = c(ab + 1)$. By assumption c is not a perfect square. By symmetry in a and b , we may assume that $a \geq b$. The quadratic equation

$$x^2 - bcx + b^2 - c = 0$$

has a solution equal to a by assumption. Let

$$a' = bc - a = \frac{b^2 - c}{a}$$

be the other solution. Note that $a' = bc - a$ is an integer and that a' is nonzero since $c \neq b^2$ (as c is not a perfect square). We claim that a' is positive. Otherwise, we would have $a' \leq -1$, thus $b^2 - c \leq -a$ and $c \geq b^2 + a$. But then

$$a^2 + b^2 = c(ab + 1) \geq (b^2 + a)(ab + 1) = ab^3 + a^2b + b^2 + a > a^2b + b^2 \geq a^2 + b^2,$$

a contradiction. Thus (a', b) is another pair satisfying the assumptions of the problem and for which the conclusion fails. By minimality of (a, b) we must have $a + b \leq a' + b$, thus $a \leq a'$. This is however impossible, since (using that $a \geq b$)

$$a' = \frac{b^2 - c}{a} < \frac{b^2}{a} \leq b.$$

Thus there are no pairs satisfying the assumptions of the problem and failing to satisfy the conclusion. \square

Example 2.37. (IMO 2007) Let a, b be positive integers such that $4ab - 1 \mid (4a^2 - 1)^2$. Prove that $a = b$.

Proof. Since $4ab \equiv 1 \pmod{4ab - 1}$, we have $4a^2b \equiv a \pmod{4ab - 1}$. Since $4ab - 1 \mid (4a^2b - b)^2$, we deduce that $4ab - 1 \mid (a - b)^2$. We argue now as in example 2.36, assuming that (a, b) is a pair satisfying $4ab - 1 \mid (a - b)^2$ and $a \neq b$, and minimizing $a + b$. We may assume that $a > b$. Write $(a - b)^2 = c(4ab - 1)$ and consider the other solution

$$a' = 2b(1 + 2c) - a = \frac{b^2 + c}{a}$$

of the equation

$$(x - b)^2 = c(4bx - 1).$$

Clearly a' is also a positive integer and (a', b) satisfies $4a'b - 1 \mid (a' - b)^2$ and $(a' - b)^2 = c(4a'b - 1)$ (thus $a' \neq b$). Using the minimality of (a, b) we

deduce that $a + b \leq a' + b$, so $a' \geq a$ and $b^2 + c \geq a^2$. However the equation $(a - b)^2 = c(4ab - 1)$ yields $c \leq (a - b)^2$, so we obtain

$$a^2 - b^2 \leq (a - b)^2.$$

Since $a > b$, this yields $a + b \leq a - b$, plainly absurd. Therefore there are no such pairs (a, b) with $a \neq b$ and the result follows. \square

Remark 2.38. Here are a few very similar problems, all of which can be solved by the same argument:

- a) Positive integers a, b satisfy $ab \mid a^2 + b^2 + 1$. Prove that $a^2 + b^2 + 1 = 3ab$.
- b) Let a, b be positive integers such that $a^2 + b^2$ is divisible by $ab - 1$. Prove that $\frac{a^2 + b^2}{ab - 1} = 5$.
- c) (AMM 11374) Let a, b, c, d be positive integers such that

$$abcd = a^2 + b^2 + c^2 + 1.$$

Prove that $d = 4$.

- d) (USA TST 2002) Find all ordered pairs of positive integers (m, n) such that $mn - 1$ divides $m^2 + n^2$.
- e) (USA TST 2009) Find all pairs of positive integers (m, n) such that $mn - 1$ divides $(n^2 - n + 1)^2$.
- f) (Hurwitz) The equation

$$x_1^2 + x_2^2 + \dots + x_n^2 = kx_1x_2\dots x_n$$

has no solutions in positive integers if $k > n$.

Example 2.39. (Kvant) Let p and q be integers greater than 1. Assume that $p \mid q^3 - 1$ and $q \mid p - 1$. Prove that $p = q^{3/2} + 1$ or $p = q^2 + q + 1$.

Proof. Write $p = qn + 1$ for some positive integer n . Then $qn + 1 \mid q^3 - 1$, so $qn + 1 \mid q^3n - n$. But

$$q^3n - n = q^2 \cdot qn - n = q^2(qn + 1) - (q^2 + n),$$

hence $qn + 1 \mid q^2 + n$. In particular $qn + 1 \leq q^2 + n$, which can be written as $n(q - 1) \leq q^2 - 1$ and yields $n \leq q + 1$.

Next, we have $qn + 1 \mid q^3n^2 - n^2$ and

$$q^3n^2 - n^2 = q^2n^2 \cdot q - n^2 = (q^2n^2 - 1)q + q - n^2.$$

Since $qn + 1$ divides $q^2n^2 - 1$, it follows that $qn + 1 \mid q - n^2$.

Now, we discuss three cases. If $q = n^2$, then $p = qn + 1 = q^{3/2} + 1$ and we are done. If $q > n^2$, then the second paragraph yields $qn + 1 \leq q - n^2 < q$, certainly impossible. Finally, if $q < n^2$, then the second paragraph yields $qn + 1 \leq n^2 - q$, thus $q(n + 1) \leq n^2 - 1$ and $q \leq n - 1$. Combined with the first paragraph, this gives $q = n - 1$ and then $p = qn + 1 = q(q + 1) + 1 = q^2 + q + 1$. The result follows. \square

Example 2.40. (Bulgaria) Let a, b and c be positive integers such that ab divides $c(c^2 - c + 1)$ and $a + b$ is divisible by $c^2 + 1$. Prove that the sets $\{a, b\}$ and $\{c, c^2 - c + 1\}$ coincide.

Proof. Write $c(c^2 - c + 1) = mab$ and $a + b = n(c^2 + 1)$ for some positive integers m, n . Without loss of generality, assume that $b \leq a$. Then

$$mab = c(c^2 - c + 1) < c(c^2 + 1) = \frac{c}{n}(a + b) \leq 2a\frac{c}{n},$$

hence $b < \frac{2c}{mn}$.

On the other hand we have $a \equiv -b \pmod{c^2 + 1}$, thus taking the equation $c(c^2 - c + 1) = mab$ modulo $c^2 + 1$ yields $1 \equiv -mb^2 \pmod{c^2 + 1}$, that is $c^2 + 1 \mid mb^2 + 1$. Thus $mb^2 + 1 = r(c^2 + 1)$ for some positive integer r . In particular $mb^2 \geq rc^2$, which combined with the inequality $b < \frac{2c}{mn}$ yields $rmn^2 < 4$. This forces $n = 1$ and $rm < 4$.

Suppose that $m > 1$, then necessarily $r = 1$. Since $r = 1$, we have $mb^2 = c^2$, in particular $b \mid c^2$. Since $b \mid mab = c(c^2 - c + 1)$ and $b \mid c^2$, we obtain $b \mid c$, thus $c = kb$ for some integer k and $k^2 = m \in \{2, 3\}$. This is clearly impossible, thus $m = 1$. It follows that the numbers a, b and $c, c^2 - c + 1$ have the same sum and product. Thus they are roots of the same quadratic polynomial and the result follows. \square

Remark 2.41. The solution would be slightly easier if we were willing to use that $\sqrt{2}$ and $\sqrt{3}$ are irrational numbers, which would immediately rule out the equation $mb^2 = c^2$ with $1 < m < 4$.

Example 2.42. (Romania TST 2012) Let a_1, \dots, a_n be positive integers and let $a > 1$ be a multiple of $a_1 a_2 \dots a_n$. Prove that $a^{n+1} + a - 1$ is not divisible by $(a + a_1 - 1)(a + a_2 - 1) \dots (a + a_n - 1)$.

Proof. Suppose that

$$a^{n+1} + a - 1 = k(a + a_1 - 1) \dots (a + a_n - 1) \quad (1)$$

for some positive integer k , and write $a = m a_1 \dots a_n$ for some positive integer m . Note that $a_1, \dots, a_n > 1$, for if $a_1 = 1$ (for example) then the right-hand side of relation (1) is divisible by a , but the left-hand side is not.

Relation (1) coupled with the congruences $a^{n+1} \equiv 1 \pmod{a - 1}$ and $a + a_i - 1 \equiv a_i \pmod{a - 1}$ for $1 \leq i \leq n$ yield

$$1 \equiv k a_1 \dots a_n \pmod{a - 1}, \quad \text{hence} \quad m \equiv k a \equiv k \pmod{a - 1}.$$

Note that $m < a = m a_1 \dots a_n$ and, since $a_i > 1$ for $1 \leq i \leq n$

$$a^{n+1} + a - 1 \geq k(a + 1)^n,$$

which easily implies that $k < a$ (since one easily checks that $a(a + 1)^n > a^{n+1} + a - 1$). Thus k, m are positive integers less than or equal to $a - 1$ and $k \equiv m \pmod{a - 1}$, which implies $k = m$. But $m \mid a$ and $k \mid a^{n+1} + a - 1$, hence $m \mid a^{n+1} + a - 1$, which implies that $m \mid 1$ and finally $k = m = 1$. It follows that

$$a^{n+1} < a^{n+1} + a - 1 = (a + a_1 - 1) \dots (a + a_n - 1),$$

which can be rewritten as

$$a_1 \cdot \dots \cdot a_n = a < \frac{a + a_1 - 1}{a} \cdot \dots \cdot \frac{a + a_n - 1}{a}.$$

This is however impossible, since for $1 \leq i \leq n$ we have

$$\frac{a + a_i - 1}{a} < a_i,$$

this inequality being equivalent to $(a - 1)(a_i - 1) > 0$.

The problem is solved. □

Example 2.43. (Schinzel) Prove that there exists a constant $c > 0$ with the following property: if a positive integer a is even and not a multiple of 10, then the sum of the digits of a^k is greater than $c \log k$ for all $k \geq 2$.

Proof. Define a sequence $(b_n)_{n \geq 0}$ by $b_0 = 0$ and $b_{n+1} = 1 + [b_n \log_2(10)]$. This sequence is increasing and $b_{n+1} \leq (1 + \log_2(10))b_n$ for $n \geq 1$, thus $b_n \leq c^n$ for all $n \geq 1$, where $c = 1 + \log_2(10)$. Suppose now that $k \geq b_n$ and write $a^k = c_0 + 10c_1 + \dots$ in base 10. For each $2 \leq j \leq n$ we have that 2^{b_j} divides a^k and since 2^{b_j} also divides $c_{b_j}10^{b_j} + c_{b_j+1}10^{b_j+1} + \dots$, it follows that 2^{b_j} divides $c_0 + 10c_1 + \dots + c_{b_j-1}10^{b_j-1}$. Note that this last number is nonzero since $c_0 \neq 0$ by assumption. We deduce that $2^{b_j} \leq c_0 + 10c_1 + \dots + c_{b_j-1}10^{b_j-1}$. Assuming that $c_{b_{j-1}}, \dots, c_{b_j-1}$ are all zero we deduce that $2^{b_j} < 10^{b_{j-1}}$, contradicting the definition of the sequence $(b_n)_{n \geq 0}$. Thus for each $2 \leq j \leq n$ there is at least one nonzero digit between $c_{b_{j-1}}, \dots, c_{b_j-1}$. It follows that if $k \geq b_n$, then the sum of digits of a^k is at least $n - 1 \geq n/2$. Taking into account that $b_n \leq c^n$ for all $n \geq 1$, the result follows. \square

2.2 Induction and binomial coefficients

The main topic of this section is the use of induction as a tool for proving divisibilities (or for solving constructive problems). Along the way, we will study some basic properties of binomial coefficients, which will help us establish a certain number of remarkable congruences. The study of binomial coefficients will occur quite frequently in this book, since they have remarkable arithmetic properties. Since we haven't developed enough theory so far, the results in this section are rather modest, but we will need them later on to obtain rather nontrivial results.

2.2.1 Proving divisibility by induction

Before studying binomial coefficients, let us spend some time dealing with examples of problems involving divisibility in which induction plays a key role.

Example 2.44. Prove that if n is a power of 3, then $n \mid 2^n + 1$.

Proof. We need to prove that 3^k divides $2^{3^k} + 1$ for all $k \geq 0$. We prove this by induction on k , the case $k = 0$ being clear. Assume that $3^k \mid 2^{3^k} + 1$ and write $2^{3^k} = n \cdot 3^k - 1$ for some integer n . Then

$$\begin{aligned} 2^{3^{k+1}} &= (2^{3^k})^3 = (n \cdot 3^k - 1)^3 \\ &= n^3 \cdot 3^{3k} - n^2 \cdot 3^{2k+1} + n \cdot 3^{k+1} - 1 \equiv -1 \pmod{3^{k+1}}, \end{aligned}$$

as needed.

We can also prove this result directly, by factoring

$$2^{3^k} + 1 = (2 + 1)(2^2 - 2 + 1)(2^{2^3} - 2^3 + 1) \dots (2^{2^{3^{k-1}}} - 2^{3^{k-1}} + 1)$$

and observing that for $i \geq 0$ we have $2^{2^{3^i}} - 2^{3^i} + 1 \equiv 0 \pmod{3}$. Hence each of the factors $2^2 - 2 + 1, 2^{2^3} - 2^3 + 1, \dots, 2^{2^{3^{k-1}}} - 2^{3^{k-1}} + 1$ is a multiple of 3. The result follows. \square

Remark 2.45. We strongly suggest the reader to try to prove by induction theorem 2.31, following the same method as the one explained in the previous example.

Example 2.46. Let n be a positive integer. Find the largest integer k for which

$$2^k \mid (n+1)(n+2) \dots (n+n).$$

Proof. Let $a_n = (n+1)(n+2) \dots (n+n)$. The first few values of the sequence $(a_n)_{n \geq 1}$ are 2, $12 = 3 \cdot 4$, $120 = 8 \cdot 15$, etc. We conjecture that the largest k for which 2^k divides a_n is n . We will prove this by induction, the case $n = 1$ being clear. In order to prove the inductive step, we will find a simple relationship between a_n and a_{n+1} . Namely,

$$\begin{aligned} a_{n+1} &= (n+2)(n+3) \dots (n+1+n+1) = (n+2) \dots (n+n)(2n+1) \cdot 2(n+1) \\ &= 2(n+1)(n+2) \dots (n+n)(2n+1) = 2a_n \cdot (2n+1). \end{aligned}$$

Since $2n+1$ is odd, the highest power of 2 dividing $2a_n(2n+1)$ is one plus the largest power of 2 dividing a_n , thus by induction this highest power is $n+1$, proving the inductive step. Hence the result of the problem is $k = n$. \square

Remark 2.47. Iterating the relation

$$a_{n+1} = 2a_n(2n+1)$$

yields the interesting equality

$$(n+1)(n+2)\dots(n+n) = a_n = 2^n \cdot 1 \cdot 3 \cdot \dots \cdot (2n-1).$$

This can also be proved directly, by observing that

$$\begin{aligned} (n+1)(n+2)\dots(n+n) &= \frac{(2n)!}{n!} = \frac{1 \cdot 3 \cdot \dots \cdot (2n-1) \cdot 2 \cdot 4 \cdot \dots \cdot 2n}{n!} \\ &= 1 \cdot 3 \cdot \dots \cdot (2n-1) \cdot \frac{2^n \cdot n!}{n!} = 2^n \cdot 1 \cdot 3 \cdot \dots \cdot (2n-1). \end{aligned}$$

Example 2.48. (IberoAmerican 2012) Let a, b, c, d be integers such that $a-b+c-d$ is an odd divisor of $a^2-b^2+c^2-d^2$. Prove that $a-b+c-d$ divides $a^n-b^n+c^n-d^n$ for all positive integers n .

Proof. By assumption $a-b+c-d$ divides $a^2-b^2+c^2-d^2$, but $a-b+c-d$ also divides $(a+c)^2-(b+d)^2$, thus it divides the difference of the two numbers, which is $2(ac-bd)$. Since $a-b+c-d$ is odd, it follows that $a-b+c-d \mid ac-bd$. We will prove by induction that $a-b+c-d$ divides $a^n-b^n+c^n-d^n$ for all positive integers n . The cases $n=1, 2$ being clear, assume that $n \geq 3$ and that $a-b+c-d \mid a^k-b^k+c^k-d^k$ for $k < n$. Let $e = a-b+c-d$. Since $a^{n-1}+c^{n-1} \equiv b^{n-1}+d^{n-1} \pmod{e}$ and $a+c \equiv b+d \pmod{e}$, we have

$$(a+c)(a^{n-1}+c^{n-1}) \equiv (b+d)(b^{n-1}+d^{n-1}) \pmod{e}.$$

Expanding and rearranging yields

$$a^n-b^n+c^n-d^n \equiv bd(b^{n-2}+d^{n-2})-ac(a^{n-2}+c^{n-2}) \equiv 0 \pmod{e},$$

the last congruence being a consequence of the congruences $bd \equiv ac \pmod{e}$ and $b^{n-2}+c^{n-2} \equiv a^{n-2}+c^{n-2} \pmod{e}$. \square

Example 2.49. Define a sequence $(a_n)_{n \geq 1}$ by setting $a_1 = 2$ and $a_{n+1} = 2^{a_n} + 2$ for $n \geq 1$. Prove that a_n divides a_{n+1} for all n .

Proof. We will prove by induction that a_n divides a_{n+1} **and** that $a_n - 1$ divides $a_{n+1} - 1$ for all $n \geq 1$. This is clear for $n = 1$, so assume that it holds for $n - 1$ (with $n \geq 2$) and let us prove it for n . Proving that $a_n \mid a_{n+1}$ reduces (thanks to the recurrence relation) to proving that $2^{a_{n-1}-1} + 1 \mid 2^{a_n-1} + 1$. For this, it suffices to check that $\frac{a_n-1}{a_{n-1}-1}$ is an odd integer. It is an integer by the inductive hypothesis, and it is clearly odd, since a_n is even for all n . Proving that $a_n - 1$ divides $a_{n+1} - 1$ reduces to $2^{a_{n-1}} + 1 \mid 2^{a_n} + 1$ and it suffices again to check that $\frac{a_n}{a_{n-1}}$ is an odd integer. The fact that it is an integer follows from the inductive hypothesis, while the fact that it is odd follows from $a_n \equiv a_{n-1} \equiv 2 \pmod{4}$ (which follows directly from the recurrence relation and the fact that $a_1 = 2$). This proves the inductive step. \square

Example 2.50. (China 2004) Prove that every positive integer n , except a finite number of them, can be represented as a sum of 2004 positive integers: $n = a_1 + a_2 + \dots + a_{2004}$, where $1 \leq a_1 < a_2 < \dots < a_{2004}$, and $a_i \mid a_{i+1}$ for all $1 \leq i \leq 2003$.

Proof. We will prove by induction on k the following statement: there exists a positive integer n_k such that all $n \geq n_k$ can be written $n = a_1 + a_2 + \dots + a_k$ for an increasing sequence $1 \leq a_1 < \dots < a_k$ with $a_1 \mid a_2 \mid \dots \mid a_k$. Call such a decomposition admissible.

The statement is trivial for $k = 1$ and for $k = 2$ we can take $n_2 = 3$ (by writing $n = 1 + (n - 1)$ for $n > 2$). Suppose now that n_k exists and choose some large n (we will make this statement more precise later on). Write $n = 2^r(2m + 1)$ for some nonnegative integers r and m . If n is large, then at least one of r and m is large.

We start with the easy case: suppose that m is large, say $m \geq n_k$. Then we can find an admissible writing $m = a_1 + a_2 + \dots + a_k$ for m and we obtain an admissible writing for n

$$n = 2^r + 2^{r+1}a_1 + 2^{r+1}a_2 + \dots + 2^{r+1}a_k.$$

Suppose now that r is large. It is enough to find an admissible decomposition for 2^r (as then we can multiply all of its members by $2m + 1$ to get an admissible writing for n). Write $r = 2q + r_1$ with $r_1 \in \{0, 1\}$ and $q \geq 0$. By the same argument, it suffices to find an admissible decomposition for 2^{2q} . Assume

that $2^q \geq n_k$ and choose an admissible decomposition $2^q + 1 = a_1 + \dots + a_k$ for $2^q + 1$. We obtain a new admissible decomposition of length $k + 1$ for 2^{2q}

$$2^{2q} = 1 + (2^q - 1)(2^q + 1) = 1 + (2^q - 1)a_1 + \dots + (2^q - 1)a_k.$$

It is now very easy to conclude: suppose that $n \geq 4n_k^3$. Then either $m \geq n_k$ or $2^q \geq n_k$. Indeed, otherwise

$$n = 2^{2q+r_1}(2m+1) < 2 \cdot n_k^2 \cdot 2n_k = 4n_k^3,$$

a contradiction. As we have explained above, this is enough to obtain an admissible decomposition of length $k + 1$ for n , so we can take $n_{k+1} = 4n_k^3$ and finish the inductive proof. \square

2.2.2 Arithmetic of binomial coefficients

We will use now induction to study binomial coefficients. Recall that if n, k are nonnegative integers with $n \geq k$, we define

$$\binom{n}{k} = \frac{n!}{k!(n-k)!},$$

where $n!$ is the product of the first n positive integers (with the convention that $0! = 1$).

A remarkable result is that $\binom{n}{k}$ is an integer (this is certainly not obvious from the definition!). There are several proofs of this result. The most standard proof consists in using a simple combinatorial argument to show that $\binom{n}{k}$ is the number of subsets with k elements of the set $\{1, 2, \dots, n\}$. We leave it to the reader to fill in the details of this combinatorial argument. Let us give an inductive proof of the fact that $\binom{n}{k}$ is an integer for all $n \geq k \geq 0$. We use strong induction on $n + k$, the cases $n + k = 0$ and $n + k = 1$ being clear. In order to prove the inductive hypothesis, we may assume that $k \geq 1$ and $n > k$ (otherwise it is clear that $\binom{n}{k} = 1$). The key point is the classical identity

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1},$$

which can be checked without any difficulty using directly the definition of $\binom{n}{k}$. Using the inductive hypothesis the numbers $\binom{n-1}{k}$ and $\binom{n-1}{k-1}$ are integers, which proves that $\binom{n}{k}$ is an integer as well. We will use this idea to prove a similar result, for which the combinatorial interpretation is not easy to find.

Example 2.51. Let q be an integer greater than 1. If n, k are nonnegative integers, define the Gaussian binomial coefficient $\binom{n}{k}_q$ by $\binom{n}{k}_q = 0$ for $k > n$ and, if $k \leq n$

$$\binom{n}{k}_q = \frac{(q^n - 1)(q^{n-1} - 1) \dots (q^{n-k+1} - 1)}{(q^k - 1)(q^{k-1} - 1) \dots (q - 1)},$$

where by convention the right-hand side equals 1 when $k = 0$.

a) Prove that for all $n, k \geq 1$ we have

$$\binom{n}{k}_q = q^k \binom{n-1}{k}_q + \binom{n-1}{k-1}_q.$$

b) Prove that $\binom{n}{k}_q$ is an integer for all n, k .

Proof. Let $x_n = q^n - 1$ for $n \geq 1$.

a) If $k > n$, then both sides are equal to 0 by definition, so assume that $k \leq n$. If $k = n$, then the equality reduces to

$$\binom{n}{n}_q = \binom{n-1}{n-1}_q$$

and holds since by definition $\binom{n}{n}_q = 1$ for all n . Finally, assume that $k \leq n-1$, then the desired equality is equivalent to

$$\frac{x_n x_{n-1} \dots x_{n-k+1}}{x_k x_{k-1} \dots x_1} = q^k \frac{x_{n-1} \dots x_{n-k}}{x_k \dots x_1} + \frac{x_{n-1} \dots x_{n-k+1}}{x_{k-1} \dots x_1}.$$

Dividing everything by $\frac{x_{n-1} \dots x_{n-k+1}}{x_{k-1} \dots x_1}$, the last relation is equivalent to

$$\frac{x_n}{x_k} = q^k \frac{x_{n-k}}{x_k} + 1$$

or $x_n = q^k x_{n-k} + x_k$. This can be checked by a direct computation.

b) This follows from part a) arguing by strong induction on $n + k$, in the same way as we did for the binomial coefficients. \square

Example 2.52. (Tournament of the Towns 2009) For each $n \geq 1$ set

$$[n]! = 1 \cdot 11 \cdot 111 \cdot \dots \cdot \underbrace{111\dots 1}_{n \text{ ones}}.$$

Prove that for all $m, n \geq 1$ the number $[n + m]!$ is divisible by $[n]![m]!$

Proof. Note that

$$[n]! = \frac{10 - 1}{9} \cdot \frac{10^2 - 1}{9} \cdot \dots \cdot \frac{10^n - 1}{9},$$

hence

$$\frac{[n + m]!}{[n]![m]!} = \frac{\prod_{i=1}^{n+m} (10^i - 1)}{\prod_{i=1}^n (10^i - 1) \cdot \prod_{i=1}^m (10^i - 1)} = \binom{n + m}{m}_{10}.$$

The result follows then from the previous example. \square

Remark 2.53. The Gaussian binomial coefficients are generalizations of the usual binomial coefficients, which correspond to the case $q = 1$. Many formulae involving binomial coefficients have analogues for Gaussian binomial coefficients. For instance, the analogue of the binomial formula (which will be discussed later on in this section) is

$$\prod_{k=0}^{n-1} (1 + q^k X) = \sum_{k=0}^n q^{\frac{k(k-1)}{2}} \binom{n}{k}_q X^k.$$

Example 2.54. Prove that $n + 1$ divides $\binom{2n}{n}$ for all positive integers n .

Proof. We have the equality

$$(n + 1) \binom{2n + 1}{n} = (2n + 1) \binom{2n}{n} = [2(n + 1) - 1] \binom{2n}{n}.$$

Taking it modulo $n + 1$ yields $\binom{2n}{n} \equiv 0 \pmod{n + 1}$. \square

Remark 2.55. The number

$$C_n = \frac{1}{n+1} \binom{2n}{n}$$

is called the n th Catalan number. These numbers have remarkable properties, for instance one can prove (not without some effort) that

$$C_{n+1} = \sum_{k=0}^n C_k C_{n-k}.$$

The Catalan numbers also appear frequently in combinatorics, for instance C_n is the number of different ways a convex polygon with $n+2$ sides can be cut into triangles by connecting vertices with straight lines (there are dozens of combinatorial interpretations of C_n !).

Example 2.56. (Romania TST 1988) Prove that for all positive integers n , the number $\prod_{k=1}^n k^{2k}$ is a multiple of $(n!)^{n+1}$.

Proof. We compute

$$\begin{aligned} \frac{1}{(n!)^{n+1}} \cdot \prod_{k=1}^n k^{2k} &= \frac{1}{n!^{n+1}} \cdot (1 \cdot 2 \cdot \dots \cdot n \cdot 2 \cdot 3 \cdot \dots \cdot n \dots \cdot n)^2 \\ &= \frac{1}{n!^{n+1}} \left(n! \cdot \frac{n!}{2!} \cdot \dots \cdot \frac{n!}{(n-1)!} \right)^2 = \frac{n!^{n-1}}{(1!2!\dots(n-1)!)^2} \\ &= \frac{n!}{1!(n-1)!} \cdot \frac{n!}{2!(n-2)!} \cdots \frac{n}{(n-1)!1!} = \prod_{k=1}^{n-1} \binom{n}{k}, \end{aligned}$$

which is clearly an integer. □

An important observation about $\binom{n}{k}$ is that

$$\binom{n}{k} = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!}$$

is a polynomial expression (of degree k) in n . This shows that one can give a meaning to $\binom{n}{k}$ even when n does not satisfy $n \geq k$ and even when n is not

an integer. More precisely, for every real (or complex) number x and every nonnegative integer k we can define

$$\binom{x}{k} := \frac{x(x-1)\dots(x-k+1)}{k!}.$$

Similarly, we can define a polynomial of degree k

$$\binom{X}{k} := \frac{X(X-1)\dots(X-k+1)}{k!}.$$

These generalized binomial coefficients still satisfy many of the usual properties of binomial coefficients, in particular the formula

$$\binom{x}{k} = \binom{x-1}{k} + \binom{x-1}{k-1}$$

still holds. Moreover we have the fundamental

Theorem 2.57. *For all $x \in \mathbf{Z}$ and all nonnegative integers k we have $\binom{x}{k} \in \mathbf{Z}$. In other words, the product of k consecutive integers is always a multiple of $k!$.*

Proof. We have already proved this result when $x \geq 0$, so assume that $x < 0$ and write $x = -y$ with $y > 0$ and integer. Then

$$\begin{aligned} \binom{x}{k} &= \frac{x(x-1)\dots(x-k+1)}{k!} = \frac{-y(-y-1)\dots(-y-k+1)}{k!} \\ &= (-1)^k \frac{y(y+1)\dots(y+k-1)}{k!} = (-1)^k \binom{y+k-1}{k}. \end{aligned}$$

Since $\binom{y+k-1}{k}$ is an integer, the theorem is proved. □

The previous theorem shows that the polynomial

$$\binom{X}{n} := \frac{X(X-1)\dots(X-n+1)}{n!}$$

takes integer values at integers. Note that $\binom{X}{n}$ does not have integer coefficients, unless $n = 1$ (its leading coefficient is $\frac{1}{n!}$, which is not an integer if $n > 1$). The following beautiful theorem describes all polynomials sending integers to integers.

Theorem 2.58. *Let f be a polynomial with rational coefficients such that $f(n)$ is an integer for any integer n . Then we can find (unique) integers a_0, a_1, \dots, a_d such that*

$$f(X) = \sum_{i=0}^d a_i \binom{X}{i}.$$

Proof. Let us first prove that for any polynomial with rational coefficients f there are rational numbers a_0, a_1, \dots, a_d (where $d = \deg f$) such that

$$f(X) = \sum_{i=0}^d a_i \binom{X}{i}.$$

We prove this by induction on $d = \deg f$, the case $d = 0$ being clear. Assuming that the result holds for polynomials of degree not exceeding $d - 1$, consider a polynomial f of degree d . Choose a_d such that $f(X) - a_d \binom{X}{d}$ has degree not exceeding $d - 1$ (namely, if a is the leading coefficient of f , choose $a_d = d!a$). By the inductive hypothesis we can write

$$f(X) - a_d \binom{X}{d} = \sum_{i=0}^{d-1} a_i \binom{X}{i}$$

for some rational numbers a_0, \dots, a_d , and thus f has the required form.

Assume now that $f(n)$ is an integer for all integers n . Then $f(0) = a_0$ is an integer, then $f(1) = a_0 + a_1$ is an integer, hence a_1 is an integer. Assuming that a_0, \dots, a_k are integers, the relation

$$f(k) = a_0 \binom{k}{0} + a_1 \binom{k}{1} + \dots + a_{k-1} \binom{k}{k-1} + a_k$$

shows that a_k is an integer. Thus a_0, \dots, a_d are actually all integers. \square

- Remark 2.59.* 1. The hypothesis that f has rational coefficients can be dropped: any polynomial with complex coefficients that sends integers to integers must have rational coefficients (we leave this as an exercise to the reader).
2. The proof of the previous theorem only used that $f(0), f(1), \dots, f(\deg f)$ are integers. In particular, it follows that if a polynomial with rational coefficients takes integer values at $\deg f + 1$ consecutive integer values, then it takes integer values at all integers.
3. As the proof shows, for any polynomial f with complex coefficients, of degree n , we can find complex numbers a_0, \dots, a_n such that

$$f(X) = \sum_{k=0}^n a_k \binom{X}{k}.$$

Moreover, the numbers a_0, \dots, a_n are unique (this follows from the last part of the proof of theorem 2.58). They are called the Mahler coefficients of f . We will see later on that much of the arithmetic properties of polynomials are captured by these coefficients (just as much of the algebraic or analytic properties of polynomials are captured by the usual coefficients).

Example 2.60. Let a_0, a_1, \dots, a_n be integers. Prove that the polynomial

$$f(X) = \sum_{k=0}^n a_k \binom{X}{k}$$

has integer coefficients if and only if $k! \mid a_k$ for $0 \leq k \leq n$.

Proof. Let $b_k = \frac{a_k}{k!}$, so that

$$f(X) = b_0 + b_1 X + b_2 X(X-1) + \dots + b_n X(X-1)\dots(X-n+1).$$

This makes it clear that if $k! \mid a_k$ for all k , then f has integer coefficients. Conversely, suppose that f has integer coefficients, then the coefficient of X^n is an integer, which means that b_n is an integer. But then

$$f(X) - b_n X(X-1)\dots(X-n+1) = b_0 + b_1 X + \dots + b_{n-1} X(X-1)\dots(X-(n-1)+1)$$

also has integer coefficients. Considering the coefficient of X^{n-1} we deduce that b_{n-1} is an integer. Continuing like this yields $b_n, b_{n-1}, \dots, b_0 \in \mathbf{Z}$, showing that $k! \mid a_k$ for all k . \square

Example 2.61. Let f be a monic polynomial of degree $n \geq 1$ with integer coefficients. Prove that if an integer d divides $f(0), f(1), \dots, f(n)$, then $d \mid n!$.

Proof. Taking into account theorem 2.58 and the remark following it, we see that $\frac{f}{d}$ is a polynomial that sends integers to integers and so can be written

$$\frac{f(X)}{d} = \sum_{k=0}^n a_k \binom{X}{k}$$

for some integers a_0, \dots, a_n . Identifying the leading coefficients on both sides we deduce that

$$\frac{1}{d} = \frac{a_n}{n!}$$

This immediately yields $d \mid n!$, as desired. \square

Example 2.62. (Putnam) Let a_1, \dots, a_n be pairwise distinct positive integers such that $a_1 a_2 \dots a_n \mid (k + a_1)(k + a_2) \dots (k + a_n)$ for all positive integers k . Prove that a_1, \dots, a_n is a permutation of $1, 2, \dots, n$.

Proof. Applying the result of the previous example to the monic polynomial

$$f(X) = (X + a_1 + 1) \dots (X + a_n + 1)$$

and to $d = a_1 a_2 \dots a_n$, we deduce that $a_1 a_2 \dots a_n \mid n!$. We may assume that $a_1 < \dots < a_n$. Then $a_1 \geq 1, a_2 \geq 2, \dots, a_n \geq n$ and since $a_1 \dots a_n \mid n!$, this forces $a_1 = 1, \dots, a_n = n$ (if one of the inequalities above was strict, then we would have $a_1 \dots a_n > n!$). The result follows. \square

2.2.3 Derivatives and finite differences

We will now make a quite interesting parallel between the usual coefficients and the Mahler coefficients of polynomials. By definition, for any polynomial P of degree n we can find numbers a_0, \dots, a_n such that

$$P(X) = \sum_{k=0}^n a_k X^k$$

and a_0, \dots, a_n are unique (these are the coefficients of P). The proof of theorem 2.58 (see the remark following theorem 2.58) also allows us to write uniquely

$$P(X) = \sum_{k=0}^n b_k \binom{X}{k}.$$

How can we characterize the numbers a_k and b_k in terms of P ? We need the following

Definition 2.63. If $P(X) = a_0 + a_1X + \dots + a_nX^n$ is a polynomial with complex coefficients, we define

- the derivative of P as the polynomial

$$P'(X) = a_1 + 2a_2X + \dots + na_nX^{n-1}.$$

In general, the k th derivative $P^{(k)}$ of P is defined by the recurrence $P^{(1)} = P'$ and $P^{(k+1)} = (P^{(k)})'$.

- the discrete derivative of P as the polynomial ΔP with

$$\Delta P(X) = P(X+1) - P(X).$$

Define $\Delta^k P$ by the recurrence relation $\Delta^1 P = \Delta P$ and

$$\Delta^{k+1} P = \Delta(\Delta^k P) = \Delta^k P(X+1) - \Delta^k P(X).$$

Let us observe that if P is a nonzero polynomial, then P' and ΔP have degree (strictly) smaller than P . Iterating this observation yields

Theorem 2.64. *For any polynomial P of degree n we have*

$$P^{(k)} = 0 = \Delta^k P \quad \text{if } k > n.$$

Let us observe that $(X^k)' = kX^{k-1}$, therefore the polynomials

$$P_k = \frac{X^k}{k!}$$

satisfy $P'_k = P_{k-1}$. Iterating this relation yields

$$P_k^{(j)} = P_{k-j} \quad \text{if } k \geq j, \quad P_k^{(j)} = 0 \quad \text{if } k < j.$$

We deduce that for the polynomial

$$P(X) = \sum_{k=0}^n a_k X^k = \sum_{k=0}^n a_k k! P_k(X)$$

we have for all $0 \leq d \leq n$

$$\frac{P^{(d)}(X)}{d!} = \sum_{k=d}^n a_k \frac{k!}{d!} P_{k-d}(X) = \sum_{k=d}^n a_k \binom{k}{d} X^{k-d}$$

and in particular

$$a_d = \frac{P^{(d)}(0)}{d!}, \quad 0 \leq d \leq n.$$

It also follows from the previous formula that if P has integer coefficients, then all coefficients of $P^{(d)}$ are multiples of $d!$ and so we obtain the interesting divisibility

$$d! \mid P^{(d)}(a) \quad \text{if } a \in \mathbf{Z}.$$

Let us study now the analogous situation for the discrete derivative. We will see that all previous results have their discrete counterparts. Consider the polynomials

$$S_k(X) = \binom{X}{k}.$$

The identity

$$\binom{X+1}{k} = \binom{X}{k} + \binom{X}{k-1}$$

is equivalent to

$$\Delta S_k = S_{k-1}.$$

We deduce that

$$\Delta^j S_k = S_{k-j} \quad \text{if } k \geq j, \quad \Delta^j S_k = 0 \quad \text{if } j > k.$$

Thus, for any polynomial

$$P(X) = \sum_{k=0}^n b_k \binom{X}{k} = \sum_{k=0}^n b_k S_k(X)$$

we have for all $0 \leq d \leq n$

$$\Delta^d P(X) = \sum_{k=0}^n b_k \Delta^d S_k(X) = \sum_{k=d}^n b_k S_{k-d}(X).$$

Recalling that $S_j(X) = \binom{X}{j}$, we obtain the analogous formula

$$\frac{\Delta^d P(X)}{d!} = \sum_{k=d}^n \frac{b_k}{k!} \binom{k}{d} X(X-1)\dots(X-(k-d+1)).$$

We are now ready to prove the

Theorem 2.65. *If P is a polynomial with complex coefficients of the form*

$$P(X) = \sum_{k=0}^n b_k \binom{X}{k},$$

then the coefficients b_0, \dots, b_n are given by

$$b_d = \Delta^d P(0)$$

for $0 \leq d \leq n$. Moreover, if P has integer coefficients, then

$$d! \mid \Delta^d P(a) \quad \text{if } a \in \mathbf{Z}.$$

Proof. For the first part, it suffices to evaluate at $X = 0$ the identity

$$\frac{\Delta^d P(X)}{d!} = \sum_{k=d}^n \frac{b_k}{k!} \binom{k}{d} X(X-1)\dots(X-(k-d+1)).$$

For the second part, note that $\frac{b_k}{k!}$ are integers for $d \leq k \leq n$ (see example 2.60). The result follows immediately by evaluating at $X = a$ the previous identity. \square

The following theorem gives a beautiful formula for $\Delta^n P$.

Theorem 2.66. *For any polynomial P and any $n \geq 1$ we have*

$$\Delta^n P(X) = \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} P(X+k).$$

Proof. We will prove this by induction on n , the case $n = 1$ being clear. Assume that the result holds for n , then

$$\begin{aligned} \Delta^{n+1} P(X) &= \Delta(\Delta^n P)(X) = \Delta^n P(X+1) - \Delta^n P(X) \\ &= \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} P(X+k+1) - \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} P(X+k) \\ &= \sum_{k=1}^{n+1} (-1)^{n+1-k} \binom{n}{k-1} P(X+k) - \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} P(X+k) \\ &= \sum_{k=0}^{n+1} (-1)^{n+1-k} P(X+k) \left(\binom{n}{k-1} + \binom{n}{k} \right) \\ &= \sum_{k=0}^{n+1} (-1)^{n+1-k} \binom{n+1}{k} P(X+k), \end{aligned}$$

as desired. \square

An immediate but very useful consequence of theorems 2.64 and 2.66 is

Corollary 2.67. *For any polynomial P and any $n > \deg P$ we have*

$$\sum_{k=0}^n (-1)^{n-k} \binom{n}{k} P(X+k) = 0.$$

Moreover, if P has integer coefficients then for all $n \geq 0$ we have

$$n! \mid \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} P(k).$$

Proof. The first statement is the combination of theorems 2.64 and 2.66. The second statement is equivalent (using theorem 2.66) to $n! \mid \Delta^n P(0)$, which follows from theorem 2.65. \square

2.2.4 The binomial formula

One of the fundamental tools used in establishing congruences is the

Theorem 2.68. (*binomial formula*) *For all complex numbers a, b and all $n \geq 1$ we have*

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k.$$

Proof. We prove this by induction, the case $n = 1$ being clear. Assume that the result holds for n , then

$$\begin{aligned} (a+b)^{n+1} &= (a+b)(a+b)^n = (a+b) \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \\ &= \sum_{k=0}^n \binom{n}{k} a^{n+1-k} b^k + \sum_{k=0}^n \binom{n}{k} a^{n-k} b^{k+1} \\ &= \sum_{k=0}^{n+1} \binom{n}{k} a^{n+1-k} b^k + \sum_{k=0}^{n+1} \binom{n}{k-1} a^{n+1-k} b^k \\ &= \sum_{k=0}^{n+1} a^{n+1-k} b^k \left(\binom{n}{k} + \binom{n}{k-1} \right) = \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k} b^k, \end{aligned}$$

as desired. \square

Explicitly, we obtain

$$(a + b)^n = a^n + na^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \dots + b^n.$$

Note that if $n \geq 2$, then all terms except the first two in the right-hand side of the previous equality are multiples of b^2 . We deduce that

$$b^2 \mid (a + b)^n - a^n - na^{n-1}b,$$

which strengthens the divisibility $b \mid (a + b)^n - a^n$. Similarly, if $n \geq 3$, we can go one step further and obtain the divisibility

$$b^3 \mid (a + b)^n - a^n - na^{n-1}b - \frac{n(n-1)}{2}a^{n-2}b^2.$$

We actually have the following fairly general congruence for polynomials with integer coefficients:

Theorem 2.69. *If P is a polynomial with integer coefficients, then for all integers a, b and all $N \geq 0$ we have*

$$P(a + b) \equiv \sum_{k=0}^N \frac{P^{(k)}(a)}{k!} b^k \pmod{b^{N+1}}.$$

In particular, for $N = 1$ this becomes

$$P(a + b) \equiv P(a) + P'(a)b \pmod{b^2}.$$

Proof. Writing P as a linear combination with integer coefficients of monomials, we reduce the proof to the case when P is a monomial, say $P(X) = X^d$ for some $d \geq 0$. Then

$$P^{(k)}(a) = d(d-1)\dots(d-k+1)a^{d-k} = \frac{d!}{(d-k)!}a^{d-k}$$

if $k \leq d$ and $P^{(k)}(a) = 0$ for $k > d$. Thus the congruence is reduced to

$$(a + b)^d \equiv \sum_{k=0}^{\min(N, d)} \binom{d}{k} a^{d-k} b^k \pmod{b^{N+1}}.$$

This is a straightforward consequence of the binomial formula. □

Example 2.70. Prove that 343 divides $2^{147} - 1$.

Proof. We have $343 = 7^3$ and $2^{147} - 1 = 8^{49} - 1 = (7 + 1)^{49} - 1$. We conclude using the binomial formula. \square

Example 2.71. Let k be an even positive integer and define a sequence $(x_n)_n$ by $x_1 = 1$ and $x_{n+1} = k^{x_n} + 1$ for $n \geq 1$.

- a) Prove that x_{n-1} divides x_n for all $n \geq 2$.
- b) Prove that x_n^2 divides $x_{n-1}x_{n+1}$ for all $n \geq 2$.

Proof. a) We prove the desired result by induction on n , the statement being clear for $n = 2$. Assume that $a = x_{n-1}$ divides $x_n = k^a + 1$, we need to prove that $k^a + 1 \mid k^{k^a+1} + 1$. Write $k^a + 1 = ab$ for some positive odd integer b . Then

$$k^{k^a+1} + 1 = k^{ab} + 1 = (ab - 1)^b + 1 = \sum_{k=0}^{b-1} (-1)^k \binom{b}{k} (ab)^{b-k},$$

the last equality being a consequence of the binomial formula and of the fact that b is odd (thus $(-1)^b + 1 = 0$). Every term in the previous sum is a multiple of ab and the result follows.

b) Let $n \geq 2$ and let $a = x_{n-1}$, so that $x_n = k^a + 1 = ab$ for some positive integer b . We need to prove that a^2b^2 divides $a(k^{ab} + 1)$. Note that a, b are odd, since k is even. But then using the binomial formula

$$a(k^{ab} + 1) = a((ab - 1)^b + 1) = a(1 + (-1)^b + ab^2(-1)^{b-1} + \dots) = a^2b^2 + \dots$$

and each term in the previous sum is a multiple of a^2b^2 . \square

Remark 2.72. A special case of the previous example is the following problem, that was proposed in a Romanian TST: prove that if n is an odd positive integer, then

$$((n-1)^n + 1)^2 \mid n(n-1)^{(n-1)^n+1} + n.$$

Example 2.73. Find a polynomial f with integer coefficients such that $27 \mid 4^n + f(n)$ for all $n \geq 1$.

Proof. Expanding using the binomial formula yields

$$4^n = (1+3)^n \equiv 1 + \binom{n}{1}3 + \binom{n}{2}3^2 \pmod{27} = 1 + 3n + \frac{9n(n-1)}{2} \pmod{27}.$$

We would like to take

$$f(x) = -\left(1 + 3x + \frac{9x(x-1)}{2}\right),$$

but the problem is that this polynomial does not have integer coefficients. This can be easily fixed, by observing for instance that

$$\frac{9n(n-1)}{2} \equiv -9n(n-1) \pmod{27}$$

for all n . We can thus choose

$$f(X) = 9X(X-1) - 1 - 3X = 9X^2 - 12X - 1. \quad \square$$

Example 2.74. (Tournament of the Towns 2011) Prove that for all $n > 1$ the number

$$1^1 + 3^3 + \cdots + (2^n - 1)^{2^n - 1}$$

is divisible by 2^n but not by 2^{n+1} .

Proof. Let

$$S_n = 1^1 + 3^3 + \cdots + (2^n - 1)^{2^n - 1}.$$

We will prove by induction on n that 2^n divides S_n and 2^{n+1} does not divide S_n . The case $n = 2$ is clear, so assume that $S_n = 2^n m$ for some odd number m . Note that

$$S_{n+1} = S_n + \sum_{k=1}^{2^n-1} (k + 2^n)^{k+2^n}.$$

The binomial formula combined with theorem 2.31 yields

$$\begin{aligned} (k + 2^n)^{k+2^n} &= (k + 2^n)^{2^n} (k + 2^n)^k \equiv (k + 2^n)^k \\ &\equiv k^k + k^{k-1} 2^n \binom{k}{1} = k^k (1 + 2^n) \pmod{2^{n+2}}. \end{aligned}$$

Thus

$$\begin{aligned} S_{n+1} &\equiv S_n + (1 + 2^n) \cdot \sum_{k=1}^{2^n-1} k^k = 2(1 + 2^{n-1})S_n \\ &= 2^{n+1}m(1 + 2^{n-1}) \equiv 2^{n+1}m \pmod{2^{n+2}}. \end{aligned}$$

Since m is odd, it follows that 2^{n+1} divides S_{n+1} but 2^{n+2} does not divide S_{n+1} , which establishes the inductive step. \square

Example 2.75. Prove that $2^n + 3^n$ is divisible by n^2 for infinitely many positive integers n .

Proof. Let n be a solution of the problem. We will look for $a > 1$ such that $n_1 = an$ is also a solution. We need to ensure that

$$a^2n^2 \mid 3^{an} + 2^{an}.$$

By assumption we can write $3^n + 2^n = bn^2$ for some positive integer n . Then, using the binomial formula, we obtain

$$3^{an} = (3^n)^a = (bn^2 - 2^n)^a = (-1)^a 2^{na} + \sum_{k=0}^{a-1} (-1)^k \binom{a}{k} 2^{nk} (bn^2)^{a-k}.$$

Choosing a odd, we need to ensure that

$$a^2n^2 \mid \sum_{k=0}^{a-1} (-1)^k \binom{a}{k} 2^{nk} (bn^2)^{a-k}$$

and the simplest way to make this happen is to impose that $a^2n^2 \mid \binom{a}{k} (bn^2)^{a-k}$ for all $0 \leq k \leq a-1$. If we choose $b = a$, the previous divisibility trivially holds for $0 \leq k \leq a-2$ (since $(bn^2)^{a-k}$ is then a multiple of $(bn^2)^2 = a^2n^4$) and it also holds for $k = a-1$ since $\binom{a}{a-1} = a$. In order to be able to choose $b = a$, we only need to check that $b > 1$ and that b is odd (which is clear as b divides $2^n + 3^n$). This reduces to $3^n + 2^n > n^2$, which follows easily by induction.

The previous discussion shows that for any solution n of the problem we can create a bigger solution. Thus it remains to check that there is at least one solution, but it is clear that 1 is a solution. \square

Remark 2.76. We will see later on that there are only two positive integers n such that $n^2 \mid 2^n + 1$, namely 1 and 3.

2.3 Euclidean division

2.3.1 The Euclidean division

In the previous sections we dealt with those properties of divisibility and congruences which follow straight from definitions. To make the theory leave the ground, we need to introduce some new ideas, and the Euclidean division is one such great idea. The following theorem lies therefore at the very heart of number theory, despite its rather simple statement and proof, since all deeper results of elementary arithmetic rely on it.

Theorem 2.77. (*Euclidean division*) *For all integers a, b with $b > 0$ there is a unique pair of integers (q, r) such that $a = bq + r$ and $0 \leq r < b$.*

Proof. Let us first prove the uniqueness of the pair. Suppose that $a = bq + r = bq_1 + r_1$, with $0 \leq r, r_1 < b$, and without loss of generality assume that $r_1 \geq r$. If $q \neq q_1$, then

$$b > r_1 - r = |r_1 - r| = |b| \cdot |q - q_1| \geq |b| = b,$$

a contradiction. Hence $q = q_1$ and $r = r_1$.

Let us now turn to the proof of the existence of (q, r) . Let q be the integer part of $\frac{a}{b}$, i.e. the largest integer not exceeding $\frac{a}{b}$. By definition, we have $q \leq \frac{a}{b} < q + 1$ and, since $b > 0$, this can be written as $0 \leq a - bq < b$. Hence we can set $r = a - bq$ and the result follows. \square

The statement and proof of the previous theorem ask for a certain number of observations, which we gather in the following series of simple but useful remarks.

Remark 2.78. a) We may be bothered by the hypothesis $b > 0$, but it is harmless, since we may always replace b by $-b$ and q by $-q$. This implies that for all integers a, b with $b \neq 0$, we can find a unique pair (q, r) with $a = bq + r$ and $0 \leq r < |b|$.

b) Uniqueness of the pair (q, r) is lost if instead of the condition $0 \leq r < |b|$ we ask for $|r| < |b|$. An example is given by $-3 = -2 \cdot 2 + 1 = -1 \cdot 2 + (-1)$.

c) Instead of choosing q the integer part of $\frac{a}{b}$, we could have chosen the integer closest to $\frac{a}{b}$. We would then obtain $|q - \frac{a}{b}| \leq \frac{1}{2}$, i.e. setting $r = a - bq$, we would have $|r| \leq \frac{b}{2}$. This can be sometimes more useful than the result of the previous theorem.

The following theorem is a simple restatement of the Euclidean division in terms of congruences. Since we will be using congruences constantly in this book, it is worth explicitly stating the result:

Theorem 2.79. *For any integers a, n with $n \neq 0$ there is a unique $0 \leq r < |n|$ such that $a \equiv r \pmod{n}$. In other words, if n is a positive integer then any integer is congruent modulo n to a unique number in the set $\{0, 1, \dots, n-1\}$.*

The numbers q, r in theorem 2.77 are called the quotient, respectively remainder of a when divided by b . Sometimes we will denote by $a \pmod{b}$ the remainder of a when divided by $b > 0$. Note that the proof of theorem 2.77 allows us to express $a \pmod{b}$ as

$$a \pmod{b} = a - b \left\lfloor \frac{a}{b} \right\rfloor.$$

In practice this is not a very convenient formula to compute $a \pmod{b}$, but it can be rather useful in more theoretical problems. Here is a classical and beautiful example:

Example 2.80. For a positive integer n , let $r(n)$ be the sum of the remainders of n when divided by $1, 2, \dots, n$. Prove that $r(n) = r(n-1)$ for infinitely many positive integers n .

Proof. Since the remainder of n when divided by k is $n - k \lfloor \frac{n}{k} \rfloor$, we have

$$r(n) = \sum_{k=1}^n \left(n - k \left\lfloor \frac{n}{k} \right\rfloor \right) = n^2 - \sum_{k=1}^n k \left\lfloor \frac{n}{k} \right\rfloor.$$

We deduce that

$$r(n) - r(n-1) = n^2 - (n-1)^2 - \sum_{k=1}^n k \left\lfloor \frac{n}{k} \right\rfloor + \sum_{k=1}^{n-1} k \left\lfloor \frac{n-1}{k} \right\rfloor.$$

Since $\lfloor \frac{n-1}{n} \rfloor = 0$, we can further write

$$r(n) - r(n-1) = 2n - 1 - \sum_{k=1}^n k \left(\left\lfloor \frac{n}{k} \right\rfloor - \left\lfloor \frac{n-1}{k} \right\rfloor \right).$$

The key observation is that $\lfloor \frac{n}{k} \rfloor - \lfloor \frac{n-1}{k} \rfloor$ is nonzero if and only if k divides n , in which case $\lfloor \frac{n}{k} \rfloor - \lfloor \frac{n-1}{k} \rfloor = 1$. This follows immediately by writing the Euclidean division $n = qk + r$ and observing that for $r \geq 1$ the Euclidean division of $n-1$ by k is simply $n-1 = qk + (r-1)$. We conclude that

$$r(n) - r(n-1) = 2n - 1 - \sum_{k|n} k.$$

We thus need to find infinitely many n such that $\sum_{k|n} k = 2n - 1$. Note that all powers of 2 have this property, since

$$\sum_{k|2^n} k = 1 + 2 + \dots + 2^{n-1} + 2^n = 2^{n+1} - 1 = 2 \cdot 2^n - 1. \quad \square$$

The most practical way to compute remainders is to use congruences combined with the following result:

Proposition 2.81. *Let a, b, n be integers with $n \neq 0$. We have $a \equiv b \pmod{n}$ if and only if a and b give the same remainder when divided by n .*

Proof. Suppose that $a \equiv b \pmod{n}$ and write $a = b + kn$ for some integer k . Let $b = qn + r$ be the Euclidean division of b by n . Then

$$a = kn + b = (k + q)n + r$$

and since $0 \leq r < |n|$ the uniqueness of the Euclidean division implies that r is also the remainder of a when divided by n . Conversely, if a and b give the same remainder r when divided by n , then n divides $a - r$ and $b - r$, thus it divides $(a - r) - (b - r) = a - b$, which gives $a \equiv b \pmod{n}$. \square

Let us see a few numerical examples showing how to use the previous proposition:

Example 2.82. Find the remainder of 73^{21} when divided by 11.

Proof. We have $73 \equiv -4 \pmod{11}$, thus

$$73^{21} \equiv (-4)^{21} = -4^{21} = -64^7 \equiv -(-2)^7 = 2^7 = 128 \equiv 7 \pmod{11},$$

thus the remainder is 7. \square

Example 2.83. Prove that for all integers n we have $n^3 \equiv 0, \pm 1 \pmod{9}$.

Proof. For all n , we have $n = 3k, 3k \pm 1$ for some integer k , by the Euclidean division. If $n = 3k$, then $(3k)^3 = 27k^3 \equiv 0 \pmod{9}$. If $n = 3k \pm 1$, then $n^3 = 27k^3 \pm 27k^2 + 9k \pm 1 \equiv \pm 1 \pmod{9}$. Hence for all n , $n^3 \equiv 0, \pm 1 \pmod{9}$. \square

Example 2.84. Consider the sequence $(a_n)_{n \geq 1}$ defined by $a_1 = 2$ and $a_{n+1} = 2^{a_n}$. Find the remainder of $a_1 + \dots + a_{254}$ when divided by 255.

Proof. We have $a_2 = 4$, $a_3 = 16$, $a_4 = 2^{16}$, etc. It is thus clear that $a_n \geq 16$ for $n \geq 3$. On the other hand, $255 = 256 - 1 = 2^8 - 1$ divides $2^k - 1$ when $8 \mid k$. Since $8 \mid a_n$ for $n \geq 3$, we have $a_{n+1} = 2^{a_n} \equiv 1 \pmod{255}$ for $n \geq 3$. Thus

$$a_1 + a_2 + \dots + a_{254} \equiv 2 + 4 + 16 + 251 \equiv 18 \pmod{255}. \quad \square$$

Example 2.85. (USAJMO 2013) Are there integers a and b such that $a^5b + 3$ and $ab^5 + 3$ are both perfect cubes?

Proof. Assume that there are such integers a, b , and write $a^5b + 3 = x^3$ and $ab^5 + 3 = y^3$. Then

$$(x^3 - 3)(y^3 - 3) = a^5b \cdot ab^5 = a^6b^6 = (ab)^6.$$

The remainders modulo 9 of any cube are 0, 1 or 8 by example 2.83. Assume that $3 \mid x$, then $3 \mid x^3 = a^5b + 3$, so $3 \mid a^5b$. Since each of a, b is congruent to 0 or ± 1 modulo 3 by the Euclidean division, we deduce that a or b is a multiple of 3. Without loss of generality, assume that $3 \mid a$, then $a^5b + 3$ is a multiple of 3 but not of 9, so cannot be a cube. Thus x, y are not multiples of 3 and so $x^3 - 3$ and $y^3 - 3$ are congruent to -2 or -4 modulo 9. Thus their product $(ab)^6$ is congruent to 4, 8, or 7 modulo 9. This is impossible, since $(ab)^6 \equiv 1 \pmod{9}$, as $(ab)^3 \equiv \pm 1 \pmod{9}$. \square

Remark 2.86. The proof of the previous example contains a proof of an important fact. If m divides $3n$ and m is not a multiple of 3, then m divides n .

2.3.2 Combinatorial arguments and complete residue systems

The fact that there are only finitely many possibilities for the remainders of integers when divided by a fixed nonzero integer n is extremely useful in practice, since it allows us to use combinatorial arguments to solve number theory problems. Among them, let us stress the following fundamental pigeonhole principle (which follows immediately from theorem 2.79).

Theorem 2.87. (*pigeonhole principle*) a) If n is a positive integer, then among any $n + 1$ integers we can find two giving the same remainder when divided by n .

b) If n is a positive integer, then among n consecutive integers there is exactly one multiple of n (and for any $0 \leq r < n$ there is exactly one congruent to r modulo n).

c) In any infinite sequence of integers we can find infinitely many terms having the same remainder when divided by n (in particular n divides the difference of any two such terms).

Let us illustrate the previous theorem with a few interesting examples.

Example 2.88. Prove that any positive integer has a multiple whose decimal representation contains the sequence 20132014.

Proof. Let n be a positive integer and choose k such that $10^k > n$. Consider the numbers $20132014 \cdot 10^k + 1, 20132014 \cdot 10^k + 2, \dots, 20132014 \cdot 10^k + n$. Each of them starts with 20132014 and one of them is a multiple of n . \square

Example 2.89. (Erdő) Prove that among n integers we can always choose some of them whose sum is a multiple of n .

Proof. Let a_1, \dots, a_n be arbitrary integers and consider the sums

$$S_k = a_1 + a_2 + \dots + a_k$$

for $1 \leq k \leq n$. If S_1, \dots, S_n give pairwise distinct remainders when divided by n , then one of these remainders is 0 and so some S_k is a multiple of n , solving the problem in this case. Otherwise, there are integers $1 \leq i < j \leq n$ such that S_i and S_j give the same remainder when divided by n . But then

$$n \mid S_j - S_i = a_{i+1} + \dots + a_j$$

and the problem is solved in this case too. \square

The next problem is a beautiful application of the previous one.

Example 2.90. (Tournament of the Towns 2002) There's a large pile of cards. On each card a number from $1, 2, \dots, n$ is written. It is known that the sum of all numbers on all of the cards is equal to $k \cdot n!$ for some k . Prove that it is possible to arrange the cards into k stacks so that the sum of numbers written on the cards in each stack is equal to $n!$.

Proof. We will argue by induction on n , the case $n = 1$ being clear. Assume that the result holds for $n - 1$. Call a card small if the number on it does not exceed $n - 1$. Let us focus only on small cards and suppose there are at least n such cards. Pick n small cards and choose a group of such cards among the n chosen cards such that the sum of the numbers on the cards of this group is a multiple of n , necessarily of the form rn for some $r \in \{1, 2, \dots, n - 1\}$. Now compress all cards in the group in a super card and label it with number r . If there are still at least n small cards after this procedure, pick again n small cards and repeat the previous procedure to create a new super card labelled with some number between 1 and $n - 1$. Repeating this process, we will end up with a certain number of super cards and at most $n - 1$ small cards. Note that the sum of the numbers on these small cards is a multiple of n , since the sum of all cards on the table was a multiple of n . Thus the sum of the numbers on the remaining small cards is of the form rn for some $r \in \{1, 2, \dots, n - 1\}$. Finally, compress these remaining small cards into a super card with label r . Now we only have cards labelled with n and a certain number of super cards labelled with $1, 2, \dots$ or $n - 1$. We can consider each card labelled with n as a super card labelled with 1, so now we have only super cards labelled with $1, 2, \dots$ or $n - 1$, and the sum of the labels on these super cards is $kn!/n = k(n - 1)!$. By

induction, we can split the super cards into k stacks with the sum of the values in each stack equal to $(n-1)!$. Since each super card is obtained by collecting some cards, it follows that the original cards can be split into k stacks such that the sum of the numbers in each stack is $n \cdot (n-1)! = n!$. The result follows. \square

Example 2.91. (Romania 1996) Let a, b, c be integers, with a even and b odd. Prove that for any positive integer n there is an integer x such that $2^n \mid ax^2 + bx + c$.

Proof. Let $f(x) = ax^2 + bx + c$. It suffices to check that $f(0), f(1), \dots, f(2^n - 1)$ give pairwise distinct remainders mod 2^n , as then among these numbers there will be a multiple of 2^n . Now, assume that $0 \leq i < j \leq 2^n - 1$ and $f(i) \equiv f(j) \pmod{2^n}$. Thus

$$2^n \mid f(j) - f(i) = a(j^2 - i^2) + b(j - i) = (j - i)(a(i + j) + b).$$

Since a is even and b is odd, $a(i + j) + b$ is odd and so necessarily $2^n \mid j - i$, contradicting the inequalities $0 < j - i < 2^n$. \square

Example 2.92. (Kvant, M 668) The sequence x_1, x_2, \dots is defined by $x_1 = 1, x_2 = 0, x_3 = 2$ and $x_{n+1} = x_{n-2} + 2x_{n-1}$ for all $n \geq 3$. Prove that for each positive integer m there are infinitely many pairs of consecutive terms of the sequence divisible by m .

Proof. Consider the terms of the sequence modulo m and denote by r_i the remainder of x_i modulo m . Note that any three consecutive terms r_i, r_{i+1}, r_{i+2} determine not only r_{i+3} but r_{i-1} too. Hence we may define r_k for nonpositive integers k and the obtained new sequence is periodic. Indeed, the number of triples of nonnegative integers less than m is not larger than m^3 and therefore there are two equal triples $(r_i, r_{i+1}, r_{i+2}) = (r_{i+a}, r_{i+a+1}, r_{i+a+2})$. Since the first triple is determined uniquely by the second one it follows that for all k we have $(r_k, r_{k+1}, r_{k+2}) = (r_{k+a}, r_{k+a+1}, r_{k+a+2})$, i.e. the sequence (r_n) is periodic. On the other hand $r_0 = x_3 - 2x_1 = 0$ and $r_{-1} = x_2 - x_0 = 0$. Hence $r_{ka-1} = r_{ka} = 0$ which shows that for all k the terms x_{ka-1} and x_{ka} of the given sequence are divisible by m . \square

Example 2.93. Prove that each integer $n > 1$ has a multiple less than n^4 whose decimal representation has at most four different digits.

Proof. Choose k such that $2^{k-1} \leq n < 2^k$. The result is easy to check when $k \leq 5$, so assume that $k \geq 6$. There are $2^k > n$ nonnegative numbers less than 10^k and having only digits 0 and 1. Two of them must give the same remainder when divided by n , hence their difference is a multiple of n . But their difference is a number with digits 0, 1, 8 or 9, which is less than $10^k < 16^{k-1} \leq n^4$ (the inequality $10^k < 16^{k-1}$ is equivalent to $1.6^k > 16$ and holds since $1.6^6 > 16$ and $k \geq 6$). \square

Another very useful observation is the following

Proposition 2.94. *Let n be a positive integer and let a_1, \dots, a_n be integers giving pairwise distinct remainders when divided by n . Then these remainders are necessarily a permutation of $0, 1, \dots, n-1$. In particular, for all $k \geq 1$ we have*

$$a_1^k + a_2^k + \dots + a_n^k \equiv 1^k + 2^k + \dots + (n-1)^k \pmod{n}.$$

Proof. This is clear. \square

Sequences a_1, \dots, a_n as in the previous proposition occur quite often in nature, for instance any sequence of n consecutive integers has this property (by theorem 2.87). Because of their importance, such sequences deserve a name:

Definition 2.95. A sequence a_1, \dots, a_n of integers is said to be a complete residue system mod n if a_1, \dots, a_n give pairwise distinct remainders when divided by n (and then the remainders of a_1, \dots, a_n must be a permutation of $0, 1, \dots, n-1$).

The following examples illustrate the concept of complete residue system.

Example 2.96. Find all positive integers n such that there exist complete residue systems a_1, a_2, \dots, a_n and b_1, b_2, \dots, b_n modulo n such that $a_1 + b_1, a_2 + b_2, \dots, a_n + b_n$ is also a complete residue system modulo n .

Proof. If n is odd, it suffices to choose any complete residue system a_1, \dots, a_n and let $b_1 = a_1, \dots, b_n = a_n$, so assume that n is even and that such $a_1, \dots, a_n, b_1, \dots, b_n$ exist. If c_1, \dots, c_n is a complete residue system, then

$$c_1 + \dots + c_n \equiv 0 + 1 + \dots + (n-1) = \frac{n(n-1)}{2} \pmod{n}.$$

Hence the hypothesis yields

$$\begin{aligned} \frac{n(n-1)}{2} + \frac{n(n-1)}{2} &\equiv \sum_{i=1}^n a_i + \sum_{i=1}^n b_i \\ &\equiv \sum_{i=1}^n (a_i + b_i) \equiv \frac{n(n-1)}{2} \pmod{n}. \end{aligned}$$

Thus n divides $\frac{n(n-1)}{2}$, which is false for n even. \square

Example 2.97. (Serbia 2012) Find all positive integers n for which one can find a permutation a_1, a_2, \dots, a_n of $1, 2, \dots, n$ such that $a_1 + 1, a_2 + 2, \dots, a_n + n$ and $a_1 - 1, a_2 - 2, \dots, a_n - n$ form complete residue systems modulo n .

Proof. Suppose that a_1, \dots, a_n is such a permutation. Then

$$1 + 2 + \dots + n \equiv (a_1 + 1) + (a_2 + 2) + \dots + (a_n + n) \pmod{n},$$

hence n divides $a_1 + \dots + a_n = \frac{n(n+1)}{2}$, and thus n is odd. Also, we have

$$2(1^2 + 2^2 + \dots + n^2) \equiv (a_1 + 1)^2 + \dots + (a_n + n)^2 + (a_1 - 1)^2 + \dots + (a_n - n)^2 \pmod{n},$$

and the last sum equals $2(a_1^2 + \dots + a_n^2 + 1^2 + \dots + n^2)$. It follows that n divides $2(1^2 + \dots + n^2) = \frac{n(n+1)(2n+1)}{3}$, hence 3 does not divide n .

Conversely, if n is odd and not divisible by 3, let a_i be the remainder of $2i$ when divided by n (with the convention that we take remainders between 1 and n , not between 0 and $n-1$). The reader can easily check that a_1, \dots, a_n satisfies all required properties (the point is that the numbers $2i$ for $1 \leq i \leq n$ give different remainders when divided by n , and so do the numbers $3i$ for $1 \leq i \leq n$). \square

Example 2.98. (Romania JBMO TST 2013) Find all positive integers $n \geq 2$ having the following property: there is a permutation $\{a_1, a_2, \dots, a_n\}$ of the set $\{1, 2, \dots, n\}$ such that the numbers $a_1 + a_2 + \dots + a_k$, where $k \in \{1, 2, 3, \dots, n\}$, form a complete residue system modulo n .

Proof. We will prove that there is such a permutation if and only if n is even. Suppose that such a permutation exists. Then n does not divide $a_1 + \dots + a_k - (a_1 + \dots + a_{k-1}) = a_k$ for $k = 2, \dots, n$. Thus we must have $a_1 = n$ and so n cannot divide $a_1 + \dots + a_n = \frac{n(n+1)}{2}$, which forces n to be even.

Conversely, assume that n is even and let $a_i = n - i + 1$ for i odd and $a_i = i - 1$ for i even. Thus the permutation is

$$n, 1, n - 2, 3, \dots$$

If $i = 2k + 1$ is odd, then

$$\begin{aligned} a_1 + a_2 + \dots + a_i &\equiv 1 + (-2) + 3 + (-4) + \dots + (2k - 1) + (-2k) \\ &\equiv -k \equiv n - \frac{i - 1}{2} \pmod{n}. \end{aligned}$$

If $i = 2k$ is even, then

$$a_1 + a_2 + \dots + a_i \equiv 1 + (-2) + \dots + (2k - 3) - (2k - 2) + 2k - 1 \equiv k \equiv \frac{i}{2} \pmod{n}.$$

It follows immediately that all partial sums $(a_1 + a_2 + \dots + a_i)_{1 \leq i \leq n}$ give distinct remainders modulo n , and the result follows. \square

We end this section with a series of miscellaneous problems in which the Euclidean division and its various consequences explained above play a crucial role.

Example 2.99. (Kvant, M 24) Let $0 < m < n$ be positive integers. Prove that there are integers $0 < q_1 < \dots < q_r$ such that $q_1 \mid q_2 \mid \dots \mid q_r$ and

$$\frac{m}{n} = \frac{1}{q_1} + \frac{1}{q_2} + \dots + \frac{1}{q_r}.$$

Proof. We use strong induction on m , the case $m = 1$ being clear. Suppose that the result holds up to $m - 1$ and let us prove it for m . Consider $n > m > 1$ and write $n = mq + r$ with $0 \leq r < m$ and $q \geq 1$. If $r = 0$, then $\frac{m}{n} = \frac{1}{q}$ and we are done. Otherwise, we have $n = m(q + 1) - (m - r)$ and

$$\frac{m}{n} = \frac{m(q + 1)}{n(q + 1)} = \frac{n + m - r}{n(q + 1)} = \frac{1}{q + 1} + \frac{m - r}{(q + 1)n}.$$

By the inductive hypothesis we can write

$$\frac{m - r}{n} = \frac{1}{q'_2} + \dots + \frac{1}{q'_r}$$

with $q'_2 \mid \dots \mid q'_r$. Letting $q_1 = q + 1$ and $q_i = (q + 1)q'_i$ for $2 \leq i \leq r$ yields the desired representation of $\frac{m}{n}$. \square

Example 2.100. We say that a positive integer n is good if the remainder of any perfect square when divided by n is a perfect square.

a) Prove that $n = 16$ is good.

b) Prove that any good number is smaller than 500.

Proof. a) Let $n = 8k + r$ be a positive integer, with $0 \leq r \leq 7$. Then $n^2 \equiv r^2 \pmod{16}$. If $r \leq 3$, the remainder of n^2 when divided by 16 is r^2 , a perfect square. If $r = 4$, the remainder is 0, while if $5 \leq r \leq 7$ the remainder is $(8 - r)^2$, again a square.

b) Suppose that $n > 500$ is good and let $q = \lfloor \sqrt{n} \rfloor$ and $r = n - q^2$. Then $0 \leq r \leq 2q$ and $q \geq 22$. Let $M = \lfloor (\sqrt{2} - 1)q \rfloor$ and finally let

$$a_k = (q + k)^2 - n.$$

It is not difficult to check that for $1 \leq k \leq M$ we have $1 \leq a_k < n$, so that a_k is the remainder of $(q + k)^2$ when divided by n . Hence we can find positive integers $b_1 < \dots < b_M$ such that $a_k = b_k^2$ for $k \leq M$. Since

$$a_M = (q + M)^2 - n \leq 2q^2 - n \leq q^2,$$

it follows that $b_M \leq q$ and so $b_k \leq q$ for all k . For $2 \leq k \leq M$ we have

$$b_k^2 - b_{k-1}^2 = a_k - a_{k-1} = 2q + 2k - 1.$$

Combined with the fact that $b_k \leq q$ this easily yields $b_k - b_{k-1} \geq 3$ (note that $b_k - b_{k-1}$ is odd by the previous relation). But then adding these inequalities yields $3(M-1) \leq b_M - b_1 \leq q-1$. Coming back to the definition of M it is not difficult to see that the last inequality is impossible for $q \geq 22$. \square

Remark 2.101. Actually the largest good number is 16, but this requires a certain number of manual computations which are not very nice.

Example 2.102. (Japan 2000) Let $n \geq 3$. Prove that there are n pairwise distinct positive integers a_1, \dots, a_n such that the product $a_1 a_2 \dots a_{i-1} a_{i+1} \dots a_n$ gives remainder 1 when divided by a_i for $1 \leq i \leq n$.

Proof. The obvious approach is to use induction, but we will see that this is slightly tricky to implement. For $n = 3$ choose the numbers 2, 3, 5. Assume that we constructed a_1, \dots, a_n and let us try to construct a_{n+1} . This should be a divisor of $a_1 \dots a_n - 1$. To make our life easy, we try the choice $a_{n+1} = a_1 \dots a_n - 1$. Unfortunately, it is no longer true that $a_1 a_3 \dots a_{n+1} \equiv 1 \pmod{a_2}$, as required. Indeed

$$a_1 a_3 \dots a_{n+1} = a_1 a_3 \dots a_n \cdot a_{n+1} \equiv 1 \cdot (-1) \equiv -1 \pmod{a_2}.$$

Since we cannot really say anything about divisors of $a_1 \dots a_n - 1$, this naive approach seems doomed.

To make things work, we start by constructing a sequence b_1, b_2, \dots, b_n such that the product of all terms except b_i gives remainder $b_i - 1$ when divided by b_i . This is fairly easy to construct since this time the previous inductive argument works: start with $b_1 = 2$ and define inductively

$$b_{n+1} = b_1 \dots b_n + 1.$$

Assuming that b_1, \dots, b_n have the property that $\prod_{j \neq i} b_j \equiv -1 \pmod{b_i}$ for $1 \leq i \leq n$, the numbers b_1, \dots, b_{n+1} have the same property, since by construction $b_{n+1} \equiv 1 \pmod{b_i}$ for $1 \leq i \leq n$.

Now, choose $a_i = b_i$ for $1 \leq i \leq n$ and $a_{n+1} = b_1 \dots b_n - 1$. Then $a_1 \dots a_n \equiv 1 \pmod{a_{n+1}}$ and moreover for $1 \leq i \leq n$ we have

$$\prod_{1 \leq j \neq i \leq n+1} a_j = \prod_{1 \leq i \neq j \leq n} b_j \cdot a_{n+1} \equiv (-1) \cdot (-1) \equiv 1 \pmod{a_i},$$

thus a_1, \dots, a_n, a_{n+1} are a solution of the problem for all $n \geq 2$. \square

Example 2.103. (St Petersburg 2013) Let a be a positive integer with 54 digits, each equal to 0 or 1. Prove that the remainder of a when divided by $33 \cdot 34 \dots 39$ is larger than 100000.

Proof. To simplify notations, let $A = 33 \cdot \dots \cdot 39$. Since a has 54 digits, each equal to 0 or 1, we can write $a = 10^{k_1} + 10^{k_2} + \dots + 10^{k_s}$ for some integers $k_1 > \dots > k_s$, with $k_1 = 53$. Write $a = Aq + r$ for the Euclidean division of a by A . The key observation is that $10^6 - 1$ divides A , as can be easily checked from

$$10^6 - 1 = (10^3 - 1)(10^3 + 1) = 9 \cdot 3 \cdot 37 \cdot 7 \cdot 11 \cdot 13.$$

Thus, $r \equiv a \pmod{10^6 - 1}$. Now, let r_1, \dots, r_s be the remainders of k_1, \dots, k_s when divided by 6. Then $10^{k_i} \equiv 10^{r_i} \pmod{10^6 - 1}$ and so

$$r \equiv a \equiv 10^{r_1} + \dots + 10^{r_s} \pmod{10^6 - 1}.$$

Note that $r_1 = 5$ as $k_1 = 53$. If

$$10^{r_1} + \dots + 10^{r_s} < 10^6 - 1,$$

the previous congruence yields

$$r \geq 10^{r_1} + \dots + 10^{r_s} > 10^5.$$

Assume that $10^{r_1} + \dots + 10^{r_s} \geq 10^6 - 1$. Since k_1, \dots, k_s are distinct numbers between 0 and 53, at most 9 of them give remainder i when divided by 6, and this holds for all $0 \leq i \leq 5$. Thus

$$10^{r_1} + \dots + 10^{r_s} \leq 9 \cdot 1 + 9 \cdot 10 + \dots + 9 \cdot 10^5 = 10^6 - 1$$

and so this inequality should be an equality, forcing $k_1 = 53$, $k_2 = 52, \dots$, $k_{54} = 0$, in other words a would have all digits equal to 1. Moreover, $r \equiv 0 \pmod{10^6 - 1}$, hence $r \geq 10^6 - 1 > 10^5$ or $r = 0$. But if $r = 0$, then A would divide $a = \frac{10^{54} - 1}{9}$, impossible since $5 \mid A$ and 5 does not divide a . The result follows. \square

2.4 Problems for practice

Basic properties

1. Prove that the last $n + 2$ digits of 5^{2^n+n+2} are the digits of 5^{n+2} , completed on the left with some zeros.
2. Is there a polynomial f with integer coefficients such that the congruence $f(x) \equiv 0 \pmod{6}$ has 2, 3 as solutions, but no other solution in the set $\{0, 1, \dots, 5\}$?
3. (Iran 2003) Is there an infinite set S such that for all distinct elements a, b of S we have $a^2 - ab + b^2 \mid a^2b^2$?
4. (Russia 2003) Is it possible to write a positive integer in every cell of an infinite chessboard in such a manner that for all integers $m, n > 100$, the sum of numbers in every $m \times n$ rectangle is divisible by $m + n$?
5. Prove that if $k > 1$ is an integer then there are infinitely many positive integers n such that $n \mid k^n + 1$.
6. (Kvant M 904) For each positive integer A with decimal representation

$$A = \overline{a_n a_{n-1} \dots a_0}$$

we set

$$F(A) = a_n + 2a_{n-1} + \dots + 2^{n-1}a_1 + 2^n a_0$$

and consider the sequence $A_0 = A, A_1 = F(A_0), A_2 = F(A_1), \dots$

- (i) Prove that there is a term A^* of this sequence such that $A^* < 20$ and $F(A^*) = A^*$.
 - (ii) Find A^* for $A = 19^{2013}$.
7. Are there infinitely many 5-tuples (a, b, c, d, e) of positive integers such that $1 < a < b < c < d < e$ and $a \mid b^2 - 1, b \mid c^2 - 1, c \mid d^2 - 1, d \mid e^2 - 1$ and $e \mid a^2 - 1$?

8. (Romania JBMO TST 2003) Let A be a finite set of positive integers with at least three elements. Prove that there are two elements of A whose sum does not divide the sum of the other elements of A .
9. (Iran 2005) Prove that there are infinitely many positive integers n such that $n \mid 3^{n+1} - 2^{n+1}$.
10. (Mathematical Reflections S 259) Let a, b, c, d, e be integers such that

$$a(b+c) + b(c+d) + c(d+e) + d(e+a) + e(a+b) = 0.$$

Prove that $a+b+c+d+e$ divides $a^5 + b^5 + c^5 + d^5 + e^5 - 5abcde$.

11. (Kazakhstan 2011) Find the smallest integer $n > 1$ such that there exist positive integers a_1, a_2, \dots, a_n for which

$$a_1^2 + \dots + a_n^2 \mid (a_1 + \dots + a_n)^2 - 1.$$

12. (Kvant 898) Find all odd integers $0 < a < b < c < d$ such that

$$ad = bc, \quad a + d = 2^k, \quad b + c = 2^m$$

for some positive integers k and m .

13. f is a polynomial with integer coefficients such that $f(n) > n$ for every positive integer n . Define a sequence $(x_n)_{n \geq 1}$ by $x_1 = 1$ and $x_{i+1} = f(x_i)$. Assuming that each positive integer has a multiple among x_1, x_2, \dots , prove that $f(X) = X + 1$.
14. (Iran 2013) Suppose that a, b are two odd positive integers such that $2ab + 1 \mid a^2 + b^2 + 1$. Prove that $a = b$.
15. (Kvant) Prove that $n^2 + 1$ divides $n!$ for infinitely many positive integers n .
16. (Vietnam 2001) Let $(a_n)_{n \geq 1}$ be an increasing sequence of positive integers such that $a_{n+1} - a_n \leq 2001$ for all n . Prove that there are infinitely many pairs (i, j) with $i < j$ such that $a_i \mid a_j$.

Induction and binomial coefficients

17. (Tournament of the Towns) Define a sequence $(a_n)_{n \geq 0}$ by $a_0 = 9$ and $a_{n+1} = a_n^3(3a_n + 4)$ for $n \geq 0$. Prove that $a_n + 1$ is a multiple of 10^{2^n} for all n .
18. Find the largest integer k which divides $8^{n+1} - 7n - 8$ for all positive integers n .
19. Let a, b be distinct integers and let n be a positive integer. Prove that $(a - b)^2 \mid a^n - b^n$ if and only if $a - b \mid nb^{n-1}$.
20. (BAMO 2012) Let n be a positive integer such that 81 divides both n and the number obtained by reversing the order of the digits of n . Prove that 81 also divides the sum of digits of n .
21. Prove that for all $n \geq 1$ the number $\frac{(2n)!(3n)!}{n!^5}$ is an integer multiple of $(n+1)^2$.
22. Find all integers a such that n^2 divides $(n+a)^n - a$ for all positive integers n .
23. (P. Erdős) Prove that every positive integer is a sum of one or more numbers of the form $2^r \cdot 3^s$, where r and s are nonnegative integers and no summand divides another.
24. (Kvant M 2274)) Let $k \geq 2$ be an integer. Find all positive integers n such that 2^k divides $1^n + 2^n + \dots + (2^k - 1)^n$.
25. Let k be an integer greater than 1 and let a_1, \dots, a_n be integers such that

$$a_1 + 2^i a_2 + 3^i a_3 + \dots + n^i a_n = 0$$

for all $i = 1, 2, \dots, k-1$. Prove that $a_1 + 2^k a_2 + \dots + n^k a_n$ is divisible by $k!$.

26. Prove that for any integer $k \geq 3$ there are k pairwise distinct positive integers such that their sum is divisible by each of the given numbers.

27. (Kvant) Prove that for any integer $n > 1$ there exist n pairwise distinct positive integers such that for any two a, b among them the number $a + b$ is divisible by $a - b$.
28. (Romania TST 1987) Let a, b, c be integers such that $a + b + c$ divides $a^2 + b^2 + c^2$. Prove that $a + b + c$ divides $a^n + b^n + c^n$ for infinitely many positive integers n .
29. (Russia 1995) Let a_1 be an integer greater than 1. Prove that there is an increasing sequence of positive integers $a_1 < a_2 < \dots$ such that

$$a_1 + a_2 + \dots + a_k \mid a_1^2 + \dots + a_k^2$$

for all $k \geq 1$.

30. Let n be a positive integer. Prove that
- a) All multiples of $10^n - 1$ which do not exceed $10^n(10^n - 1)$ have sum of digits $9n$.
 - b) The sum of digits of any multiple of $10^n - 1$ is at least $9n$.
31. (USAMO 1998) Prove that for each $n \geq 2$ there is a set S of n integers such that $(a - b)^2$ divides ab for every distinct $a, b \in S$.
32. (Romania JBMO TST 2004) Let A be a set of positive integers such that
- a) if $a \in A$, then all positive divisors of a are also in A ;
 - b) if $a, b \in A$ satisfy $1 < a < b$, then $1 + ab \in A$.
- Prove that if A has at least 3 elements, then A is the set of all positive integers.
33. (USAMO 2002) Let a, b be integers greater than 2. Prove that there exists a positive integer k and a finite sequence n_1, n_2, \dots, n_k of positive integers such that $n_1 = a$, $n_k = b$, and $n_i n_{i+1}$ is divisible by $n_i + n_{i+1}$ for each i ($1 \leq i < k$).
34. Is it true that for any integer $k > 1$ we can find an integer $n > 1$ such that k divides each of the numbers $\binom{n}{1}, \binom{n}{2}, \dots, \binom{n}{n-1}$?

35. (Catalan) Prove that $m!n!(m+n)!$ divides $(2m)!(2n)!$ for all positive integers m, n .
36. Let $x_1 < x_2 < \dots < x_{n-1}$ be consecutive positive integers such that $x_k \mid k \binom{n}{k}$ for all $1 \leq k \leq n-1$. Prove that x_1 equals 1 or 2.

Euclidean division

37. Prove that for any $n > 1$ there are $2n-2$ positive integers such that the average of any n of them is not an integer.
38. Let n be a positive integer. Find the remainder of 3^{2^n} when divided by 2^{n+3} .
39. (Saint Petersburg 1996) Let P be a polynomial with integer coefficients, of degree greater than 1. Prove that there is an infinite arithmetic progression none of whose terms belongs to $\{P(n) \mid n \in \mathbf{Z}\}$.
40. (Baltic Way 2011) Determine all positive integers d such that whenever d divides a positive integer n , d also divides any integer obtained by rearranging the digits of n .
41. (Russia) A convex polygon on the coordinate plane contains at least $m^2 + 1$ points with integer coordinates in its interior. Show that some $m+1$ of these points lie on a line.
42. (IMO 2001) Let $n > 1$ be an odd integer and let c_1, c_2, \dots, c_n be integers. For each permutation $a = a_1, a_2, \dots, a_n$ of $1, 2, \dots, n$, define

$$S(a) = c_1 a_1 + c_2 a_2 + \dots + c_n a_n.$$

Prove that there are permutations $a \neq b$ of $1, 2, \dots, n$ such that $n! \mid S(a) - S(b)$.

43. Let $n, k > 1$ be integers. Consider a set A of k integers. For each nonempty subset B of A , compute the remainder of the sum of elements of B when divided by n . Assume that 0 does not appear among these remainders. Prove that there are at least k distinct remainders obtained

in this way. Moreover, if there are only k such remainders, then all elements of A give the same remainder when divided by n .

44. (IMO 2005) A sequence a_1, a_2, \dots of integers has the following properties:

a) a_1, a_2, \dots, a_n is a complete residue system modulo n for all $n \geq 1$.

b) there are infinitely many positive and infinitely many negative terms in the sequence.

Prove that each integer appears exactly once in this sequence.

45. For a positive integer n , consider the set

$$S = \{0, 1, 1 + 2, 1 + 2 + 3, \dots, 1 + 2 + 3 + \dots + (n - 1)\}$$

Prove S is a complete residue system modulo n if and only if n is a power of 2.

46. (Argentina 2008) 101 positive integers are written on a line. Prove that we can write signs $+$, signs \times and parentheses between them, without changing the order of the numbers, in such a way that the resulting expression makes sense and the result is divisible by $16!$.

47. (adapted from Kvant M33) Consider the remainders of 2^n when divided by $1, 2, \dots, n$. Prove that their sum exceeds $cn \log n$ for some constant $c > 0$ (independent of $n > 1$).

Chapter 3

GCD and LCM

This relatively short chapter discusses properties of the greatest common divisor and of the least common multiple of several integers, with special emphasis on the applications of these concepts to diophantine equations. Key results proved and discussed at length in this chapter are Bézout's theorem and Gauss' lemma. These are crucial results in arithmetic, which will appear constantly throughout this book.

3.1 Bézout's theorem and Gauss' lemma

3.1.1 Bézout's theorem and the Euclidean algorithm

In this chapter we will be interested in common divisors of two or several integers. We start by introducing the key definition and notation for this notion:

Definition 3.1. Let a_1, a_2, \dots, a_n be integers, not all equal to 0. We denote by $\gcd(a_1, a_2, \dots, a_n)$ and call the greatest common divisor of a_1, \dots, a_n the largest positive integer that divides a_1, a_2, \dots, a_n simultaneously.

The fact that the previous definition makes sense deserves an explanation: we need to check that the set of positive common divisors of a_1, \dots, a_n has a greatest element. This set is nonempty, since it contains 1, and this set is finite,

since any common divisor of a_1, \dots, a_n does not exceed $\max(|a_1|, |a_2|, \dots, |a_n|)$ (if all a_i 's are nonzero, we can replace $\max(|a_1|, \dots, |a_n|)$ with $\min(|a_1|, \dots, |a_n|)$), so there are only finitely many common divisors. Note that this crucially uses the hypothesis that a_1, \dots, a_n are not simultaneously equal to 0. We will take the convention that $\gcd(a_1, \dots, a_n) = 0$ when $a_1 = \dots = a_n = 0$.

By definition, $\gcd(a_1, \dots, a_n)$ divides a_1, \dots, a_n , hence it divides any linear combination of a_1, \dots, a_n . The fundamental result in this section states that $\gcd(a_1, \dots, a_n)$ is actually **equal** to some linear combination of a_1, \dots, a_n . The Euclidean division plays a crucial role in the proof.

Theorem 3.2. (*Bézout*) *For any integers a_1, \dots, a_n there are integers x_1, \dots, x_n such that*

$$\gcd(a_1, \dots, a_n) = a_1x_1 + \dots + a_nx_n.$$

Proof. If $a_1 = \dots = a_n = 0$, choose $x_1 = \dots = x_n = 0$, so assume that not all a_i 's are equal to 0. Let S be the set of all linear combinations $a_1x_1 + \dots + a_nx_n$ with integer coefficients x_1, \dots, x_n . Note that $a_1^2 + \dots + a_n^2$ is a positive integer in S , so there is a smallest positive integer d in S . We will prove that $d = \gcd(a_1, \dots, a_n)$, which implies the desired result.

Since d is a linear combination of a_1, \dots, a_n , d is a multiple of $\gcd(a_1, \dots, a_n)$. It suffices therefore to prove that d divides a_1, \dots, a_n in order to conclude. We will prove that d divides any element s of S , and in particular it divides a_1, \dots, a_n . Let $s \in S$ and suppose that d does not divide s . Thus $s = qd + r$ for some integers r, s with $0 < r < d$. Now s and d are linear combinations of a_1, \dots, a_n , thus $r = s - qd$ is also a linear combination of a_1, \dots, a_n and so r is a positive element of S smaller than d . This contradicts the minimality of d and finishes the proof of the theorem. \square

We record the following simple consequence of theorem 3.2, which will be constantly used from now on.

Corollary 3.3. *If x_1, \dots, x_n are integers and a is a positive integer, then*

$$\gcd(ax_1, \dots, ax_n) = a \cdot \gcd(x_1, \dots, x_n).$$

Proof. The result is clear if $x_1 = \dots = x_n = 0$, so assume that this is not the case. Let $d = \gcd(ax_1, \dots, ax_n)$ and $e = \gcd(x_1, \dots, x_n)$. Since $e \mid x_i$ for all i ,

we have $ae \mid ax_i$ for all i , hence ae is a common positive divisor of ax_1, \dots, ax_n and so $ae \leq d$. By the previous theorem e is a linear combination of x_1, \dots, x_n , hence ae is a linear combination of ax_1, \dots, ax_n and so ae is a multiple of d . It follows that $ae = d$, as needed. \square

Example 3.4. (Putnam 2000) Prove that the expression $\frac{\gcd(m,n)}{n} \binom{n}{m}$ is an integer for all pairs of integers $n \geq m \geq 1$.

Proof. Write $\gcd(m, n) = an + bm$ for some integers a, b , then

$$\frac{\gcd(m, n)}{n} \binom{n}{m} = a \binom{n}{m} + b \frac{m}{n} \binom{n}{m},$$

thus it suffices to check that $\frac{m}{n} \binom{n}{m}$ is an integer. But

$$\frac{m}{n} \binom{n}{m} = \frac{m}{n} \cdot \frac{n!}{(n-m)!m!} = \frac{(n-1)!}{(m-1)!(n-m)!} = \binom{n-1}{m-1}$$

is an integer. \square

We will try to find a practical way of computing $\gcd(a_1, \dots, a_n)$. The obvious and naive approach consists in testing whether k divides a_1, \dots, a_n for $1 \leq k \leq \max(|a_1|, \dots, |a_n|)$ (if all a_i 's are nonzero, we can replace \max with \min) and take the largest such k . This is not efficient at all.

We will first simplify the problem by reducing it to the case $n = 2$. In order to do this, we need the following very important result, which is an easy consequence of theorem 3.2, but which would not be so easy to prove directly from the definition of $\gcd(a_1, \dots, a_n)$.

Corollary 3.5. *Let a_1, \dots, a_n be integers. Any common divisor of a_1, \dots, a_n divides $\gcd(a_1, \dots, a_n)$.*

Proof. Any common divisor of a_1, \dots, a_n divides any linear combination of a_1, \dots, a_n and, by theorem 3.2, $\gcd(a_1, \dots, a_n)$ is a linear combination of a_1, \dots, a_n . \square

The previous corollary easily implies the following property of gcd, which reduces the computation of the gcd of n numbers to that of the gcd of $n - 1$ numbers and the gcd of two numbers. Inductively, this reduces therefore the problem of computing the gcd of n numbers to that of computing the gcd of two numbers.

Theorem 3.6. *For all integers a_1, \dots, a_n we have*

$$\gcd(a_1, \dots, a_n) = \gcd(\gcd(a_1, \dots, a_{n-1}), a_n).$$

Proof. Let $d = \gcd(a_1, \dots, a_n)$ and $e = \gcd(a_1, \dots, a_{n-1})$. Note that d is a common divisor of a_1, \dots, a_{n-1} , thus $d \mid e$ thanks to the previous corollary. We need to check that $\gcd(e, a_n) = d$. Since d divides a_n and e , we know that $d \leq \gcd(e, a_n)$. On the other hand, $\gcd(e, a_n)$ divides e and a_n , thus it divides a_1, \dots, a_n and thanks to the previous corollary again, $\gcd(e, a_n) \mid d$, thus $\gcd(e, a_n) \leq d$. We conclude that $\gcd(e, a_n) = d$. \square

The formal reductions of the problem being done, we need to solve the problem of computing $\gcd(a, b)$ for two integers a, b . The key observation is the following:

Proposition 3.7. *Let a, b be integers with $b \neq 0$ and let $a = bq + r$ be the Euclidean division of a by b . Then $\gcd(a, b) = \gcd(b, r)$.*

Proof. Any common divisor of a and b divides $a - bq = r$ and so is a common divisor of b and r . Conversely, any common divisor of b and $r = a - bq$ is a divisor of a and so a common divisor of a and b . The result follows then straight from the definition of $\gcd(a, b)$ and $\gcd(b, r)$. \square

Using the previous proposition, we obtain a very efficient way of computing $\gcd(a, b)$. If $a = 0$, then clearly $\gcd(a, b) = |b|$, and if $b = 0$ then $\gcd(a, b) = |a|$. Thus we may assume that $a, b \neq 0$. Also, $\gcd(a, b) = \gcd(|a|, |b|)$, so replacing a and b with their absolute values we may assume that a, b are positive integers. Finally, $\gcd(a, b) = \gcd(b, a)$, so we may assume that $a \geq b$. Then we apply

the Euclidean division and obtain the relations

$$\begin{aligned}
 a &= bq_1 + r_1, & 0 \leq r_1 < b \\
 b &= r_1q_2 + r_2, & 0 \leq r_2 < r_1 \\
 r_1 &= r_2q_3 + r_3, & 0 \leq r_3 < r_2 \\
 &\dots & \dots \\
 r_{k-2} &= r_{k-1}q_k + r_k, & 0 \leq r_k < r_{k-1} \\
 r_{k-1} &= r_kq_{k+1} + r_{k+1}, & r_{k+1} = 0
 \end{aligned}$$

Since $b > r_1 > r_2 > \dots$ are nonnegative integers, there must be some k for which $r_{k+1} = 0$. Hence our process must terminate. Moreover, by the previous proposition

$$d = \gcd(a, b) = \gcd(b, r_1) = \gcd(r_1, r_2) = \dots = \gcd(r_k, r_{k+1}) = \gcd(r_k, 0) = r_k,$$

thus $\gcd(a, b)$ is the last nonzero remainder obtained in the process. We have just proved the very important

Theorem 3.8. (*Euclidean algorithm*) Let $a > b$ be positive integers. Define $r_0 = a$, $r_1 = b$ and, as long as $r_n \neq 0$, define r_{n+1} as the remainder in the Euclidean division of r_{n-1} by r_n . Then there is a smallest $n \geq 1$ for which $r_n = 0$, and $r_{n-1} = \gcd(a, b)$.

Example 3.9. Compute

- a) $\gcd(2050, 123)$.
- b) $\gcd(987654321, 123456789)$.
- c) $\gcd(2016, 2352, 1680)$.

Proof. a) The Euclidean algorithm is implemented as follows

$$\begin{aligned}
 2050 &= 123 \cdot 16 + 82 \\
 123 &= 82 \cdot 1 + 41 \\
 82 &= 2 \cdot 41 + 0
 \end{aligned}$$

Hence $\gcd(2050, 123) = 41$.

b) Set $a = 987654321$, $b = 123456789$. Euclidean division yields $a = 8b + 9$. Next, we need to perform the Euclidean division of b by 9. A direct

computation shows that $9 \mid b$ and so the corresponding remainder is 9. It follows that $\gcd(a, b) = 9$.

c) We first find

$$\gcd(1680, 2016) = \gcd(16 \cdot 105, 16 \cdot 126) = 16 \gcd(105, 126) = 16 \cdot 21.$$

We next find

$$\gcd(16 \cdot 21, 2352) = \gcd(16 \cdot 21, 16 \cdot 147) = 16 \cdot \gcd(21, 147) = 16 \cdot 21 = 336.$$

Thus the answer is 336. \square

3.1.2 Relatively prime numbers

We move on to the second important topic of this section, that of coprime and pairwise relatively prime numbers. Let us define this concept first.

Definition 3.10. Integers a_1, \dots, a_n are called relatively prime or coprime if

$$\gcd(a_1, \dots, a_n) = 1.$$

They are called pairwise relatively prime if $\gcd(a_i, a_j) = 1$ for all $1 \leq i \neq j \leq n$.

Remark 3.11. Saying that a_1, \dots, a_n are pairwise relatively prime is much stronger than saying that a_1, \dots, a_n are relatively prime. For instance 6, 10, 15 are coprime since no integer greater than 1 divides all of them, but $\gcd(6, 10) = 2 > 1$, $\gcd(6, 15) = 3 > 1$ and $\gcd(10, 15) = 5 > 1$.

Before moving on to more technical things, let us give some classical examples illustrating the previous notions. The following example is very important, establishing a key property of the numbers

$$F_n = 2^{2^n} + 1,$$

called Fermat's numbers. These numbers play a fundamental role in arithmetic and quite a lot of difficult problems concern them (we will see the appearance of Fermat numbers quite often in this book). The following problem shows that these numbers are pairwise relatively prime (note that it is not entirely obvious how to construct infinite sequences of positive integers such that any two terms in the sequence are relatively prime).

Example 3.12. Let $F_n = 2^{2^n} + 1$ be the n th Fermat number. Prove that

$$\gcd(F_m, F_n) = 1$$

for $m \neq n$.

Proof. We may assume that $m > n$. Suppose that $d > 1$ is a common divisor of F_m and F_n , then clearly d is odd, since F_n is odd. Since $2^{2^n} \equiv -1 \pmod{d}$, we also have

$$(2^{2^n})^{2^{m-n}} \equiv (-1)^{2^{m-n}} \equiv 1 \pmod{d},$$

in other words $2^{2^m} \equiv 1 \pmod{d}$. But by assumption $d \mid F_m$, thus $2^{2^m} \equiv -1 \pmod{d}$. We deduce that $d \mid 2$ and since d is odd, we must have $d = 1$. \square

An alternative argument which can be used to prove that the Fermat numbers are pairwise relatively prime is based on the identity

$$F_n - 2 = F_0 F_2 \dots F_{n-1},$$

which follows from

$$2^{2^n} - 1 = (2 - 1)(2 + 1)(2^2 + 1) \dots (2^{2^{n-1}} + 1).$$

Thus if d divides F_n and F_m with $m < n$, then d divides $2 = F_n - F_0 \dots F_{n-1}$. Since d is odd, we must have $d = 1$.

The next example is a variation on this theme.

Example 3.13. Let f be a polynomial with integer coefficients such that $f(0) = f(1) = 1$. Prove that for all integers n , the numbers $n, f(n), f(f(n)), f(f(f(n))), \dots$ are pairwise relatively prime.

Proof. Let n be an integer and define the sequence $(a_k)_{k \geq 0}$ by $a_0 = n$ and $a_{k+1} = f(a_k)$ for $k \geq 0$. We need to prove that a_0, a_1, \dots are pairwise relatively prime.

By hypothesis $f - 1$ vanishes at 0 and 1, thus we can write

$$f(X) = X(X - 1)g(X) + 1$$

for some polynomial g with integer coefficients. Then

$$a_{k+1} = f(a_k) = 1 + a_k(a_k - 1)g(a_k),$$

which can be written as

$$a_{k+1} - 1 = (a_k - 1)a_k g(a_k).$$

A straightforward induction, then gives

$$a_m - 1 = (a_0 - 1) \prod_{k=0}^{m-1} (a_k g(a_k)).$$

The right-hand side is a multiple of $a_0 a_1 \dots a_{m-1}$. Thus if d divides a_j for some $j < m$, then d divides $a_m - 1$ and d does not divide a_m unless $d = 1$. We deduce that $\gcd(a_j, a_k) = 1$ for $j < k$ and the result follows. \square

Example 3.14. (Miklos Schweitzer Competition 1949) Let n and k be positive integers, $n \geq k$. Prove that the numbers $\binom{n}{k}, \binom{n+1}{k}, \dots, \binom{n+k}{k}$ are relatively prime.

Proof. We prove this by induction on k , the case $k = 1$ being clear. Assume that the result holds for $k - 1$ and let d be a common divisor of $\binom{n}{k}, \binom{n+1}{k}, \dots, \binom{n+k}{k}$. Then d divides the successive differences between these numbers, thus $d \mid \binom{n+1}{k} - \binom{n}{k} = \binom{n}{k-1}$, then $d \mid \binom{n+2}{k} - \binom{n+1}{k} = \binom{n+1}{k-1}$ and finally $d \mid \binom{n+k-1}{k-1}$. But by the inductive hypothesis the numbers $\binom{n}{k-1}, \dots, \binom{n+k-1}{k-1}$ are relatively prime, so $d \mid 1$ and the result follows. \square

Example 3.15. (Tournament of the Towns 2003) An increasing arithmetic progression consists of one hundred positive integers. Is it possible that every two of them are relatively prime?

Proof. Yes, it is possible. We are looking for positive integers a, b such that $a + ib$ and $a + jb$ are relatively prime for $0 \leq i < j \leq 99$. Suppose that d divides $a + ib$ and $a + jb$, then it divides $(j - i)b$ and so it divides $99!b$. But since $d \mid 99!a + i99!b$, d also divides $99!a$, and so $d \mid 99! \gcd(a, b)$. We choose a, b relatively prime, then $d \mid 99!$. Next, we choose b a multiple of $99!$, then d divides a (since it divides $99!$ and $a + ib$). Finally, choosing $a = 1$ (or any number congruent to 1 modulo $99!$) yields the desired arithmetic progression. \square

Example 3.16. (Kvant, M 1014) Let a_1, a_2, \dots, a_n be pairwise distinct and pairwise relatively prime numbers. Prove that there are infinitely many positive integers b such that the numbers $a_1 + b, a_2 + b, \dots, a_n + b$ are also pairwise relatively prime.

Proof. Denote by P the absolute value of the product of all numbers $a_i - a_j$, $1 \leq i < j \leq n$. Then for each positive integer k the numbers $a_1 + kP, a_2 + kP, \dots, a_n + kP$ are pairwise relatively prime. Indeed, let d be a common divisor of $a_i + kP$ and $a_j + kP$. Then d divides $a_i - a_j$ and hence it divides P . Hence d divides both a_i and a_j , i.e. $d = 1$. \square

The following result (which will be constantly used from now on) explains why relatively prime integers are simultaneously a natural and useful notion:

Proposition 3.17. *Let a_1, \dots, a_n be integers and let $d = \gcd(a_1, \dots, a_n)$. There are relatively prime integers x_1, \dots, x_n such that $a_i = dx_i$ for $1 \leq i \leq n$.*

Proof. Since d divides a_1, \dots, a_n , we can write $a_i = dx_i$ for some integers x_1, \dots, x_n . If $d = 0$, we have $a_1 = \dots = a_n = 0$ and we can take $x_i = 1$ for $1 \leq i \leq n$. If $d \neq 0$, then x_1, \dots, x_n are relatively prime, since if $e > 1$ is a common divisor of x_1, \dots, x_n , then ed is a common divisor of a_1, \dots, a_n and $ed > d$, a contradiction. \square

Theorem 3.2 yields the following characterization of relatively prime numbers:

Corollary 3.18. *Integers a_1, \dots, a_n are relatively prime if and only if there are integers x_1, \dots, x_n such that $a_1x_1 + \dots + a_nx_n = 1$.*

Proof. If there are such integers x_1, \dots, x_n , then clearly any common divisor of a_1, \dots, a_n divides $1 = a_1x_1 + \dots + a_nx_n$ and so $\gcd(a_1, \dots, a_n) = 1$. The converse follows directly from theorem 3.2. \square

We can give a slight improvement of the previous corollary (for $n = 2$) in which we take care of positivity:

Corollary 3.19. *If a, b are relatively prime positive integers, then we can find positive integers m, n such that $am - bn = 1$.*

Proof. Choose $x, y \in \mathbf{Z}$ such that $ax + by = 1$. For all integers t we have $a(x + bt) - b(at - y) = 1$, hence it is enough to show that we can find t such that $x + bt$ and $at - y$ are positive integers. Simply choose $t > \max(-x, y)$. \square

3.1.3 Inverse modulo n and Gauss' lemma

The first part of the following fundamental theorem follows straight from theorem 3.2.

Theorem 3.20. *If $\gcd(a, b) = 1$, then we can find an integer x such that $ax \equiv 1 \pmod{b}$. Moreover, any two such integers x are congruent modulo b .*

Proof. As we have already observed, only the second statement needs a proof. If x, x' are two such integers then $ax \equiv 1 \equiv ax' \pmod{b}$ and so

$$x' \equiv axx' = (ax')x \equiv x \pmod{b},$$

as needed. \square

Remark 3.21. 1) The converse of the previous result also holds, for if $ax \equiv 1 \pmod{b}$, then we can write $ax - 1 = by$ for some integer y , hence any common divisor of a and b will divide 1.

2) By the theorem, all numbers x satisfying $ax \equiv 1 \pmod{b}$ give the same remainder when divided by b . This remainder is called the inverse of a modulo b and denoted $a^{-1} \pmod{b}$.

The previous theorem has many important consequences, which wouldn't be easy to prove directly. For instance, it immediately implies the following result, which is of utmost importance and will be used throughout the book:

Theorem 3.22. (*Gauss' lemma*) *If a, b, c are integers such that $a \mid bc$ and $\gcd(a, b) = 1$, then $a \mid c$.*

Proof. Let x be an integer such that $bx \equiv 1 \pmod{a}$ (such x exists by theorem 3.20). Since $bc \equiv 0 \pmod{a}$ we obtain $xbc \equiv 0 \pmod{a}$ and so $c \equiv 0 \pmod{a}$. The result follows. \square

Let us write Gauss' lemma in terms of congruences:

Corollary 3.23. *If $ab \equiv ac \pmod{n}$, then $b \equiv c \pmod{\frac{n}{\gcd(a,n)}}$. In particular, if $\gcd(a,n) = 1$, then $b \equiv c \pmod{n}$.*

Proof. Let $d = \gcd(a,n)$ and write $a = du, n = dv$, with $\gcd(u,v) = 1$. Then $ab \equiv ac \pmod{n}$ is equivalent to $v \mid u(b - c)$. By Gauss' lemma, this is equivalent to $v \mid b - c$, i.e. $b \equiv c \pmod{v}$. \square

Another very important result is the following direct consequence of Gauss' lemma.

Theorem 3.24. *Let a, b, c be integers such that $a \mid c$, $b \mid c$ and $\gcd(a,b) = 1$. Then $ab \mid c$. In other words, if an integer is a multiple of two relatively prime numbers, then it is a multiple of their product.*

Proof. We can write $c = ad$ for some integer d . Since $b \mid c$ and $\gcd(a,b) = 1$, by Gauss' lemma we have $b \mid d$. Thus $ab \mid ad = c$ and we are done. \square

Remark 3.25. An immediate induction shows that if an integer is a common multiple of finitely many pairwise relatively prime integers, then it is a multiple of their product.

We also mention the following very useful consequence of theorem 3.20.

Corollary 3.26. *If an integer a is relatively prime to each of the integers b_1, b_2, \dots, b_n , then it is also relatively prime to $b_1 b_2 \dots b_n$.*

Proof. By theorem 3.20 we can find integers x_i such that $b_i x_i \equiv 1 \pmod{a}$. Then $(b_1 b_2 \dots b_n) \cdot (x_1 \dots x_n) \equiv 1 \pmod{a}$, hence a and $b_1 \dots b_n$ are relatively prime. \square

The following result would be fairly difficult to prove using only formal properties of the divisibility relation:

Corollary 3.27. *If a, b are integers and $a^n \mid b^n$ for some $n \geq 1$, then $a \mid b$.*

Proof. If $a = 0$ or $b = 0$, then the result is immediate, so assume that a, b are nonzero. Let $d = \gcd(a,b)$ and write $a = du$ and $b = dv$ for some relatively prime integers u, v . Then $d^n u^n \mid d^n v^n$, so $u^n \mid v^n$. By the previous corollary $\gcd(u, v^n) = 1$, and since u divides v^n (as $u \mid u^n \mid v^n$), we deduce that $u \mid 1$ and so $u = \pm 1$. Thus $a = \pm d$ and it clearly divides $b = dv$. \square

Let us see how the previous theoretical results work in practice.

Example 3.28. (Saint Petersburg) Find all relatively prime positive integers x, y such that

$$2(x^3 - x) = y^3 - y.$$

Proof. Write the equation as

$$2x^3 - y^3 = 2x - y$$

and let $z = 2x - y$. Since $\gcd(x, y) = 1$ we have $\gcd(x, z) = 1$. Next, $z \mid 2x - y$ hence $z \mid 8x^3 - y^3$ and by hypothesis $z \mid 2x^3 - y^3$, thus $z \mid 6x^3$. Since $\gcd(x, z) = 1$, we have $\gcd(z, x^3) = 1$ (corollary 3.26) and Gauss' lemma yields $z \mid 6$. We deduce that $z \in \{-1, -2, -3, -6, 1, 2, 3, 6\}$. Solving in each case the corresponding system

$$2x - y = z, \quad 2x^3 - y^3 = z$$

yields the solutions $(x, y) \in \{(1, 1), (4, 5)\}$. □

Example 3.29. (Erdős-Szekeres) Let n be a positive integer, and let k and m be positive integers such that $0 < m \leq k < n$. Prove that the numbers $\binom{n}{k}$ and $\binom{n}{m}$ are not relatively prime.

Proof. Assume that the numbers $\binom{n}{k}$ and $\binom{n}{m}$ are relatively prime and note that

$$\begin{aligned} \binom{n}{k} \cdot \binom{k}{m} &= \frac{n!}{k! (n-k)!} \cdot \frac{k!}{m! (k-m)!} \\ &= \frac{n!}{m! (n-m)!} \cdot \frac{(n-m)!}{(k-m)! (n-k)!} = \binom{n}{m} \cdot \binom{n-m}{k-m}. \end{aligned}$$

Thus $\binom{n}{k} \cdot \binom{k}{m}$ is divisible by $\binom{n}{m}$ and since $\binom{n}{k}$ and $\binom{n}{m}$ are relatively prime, it follows that $\binom{k}{m}$ is divisible by $\binom{n}{m}$. Thus $\binom{k}{m} \geq \binom{n}{m}$ (note that $\binom{k}{m} \neq 0$, as $0 < m \leq k$). This is impossible, since by hypothesis $k < n$, thus

$$\binom{k}{m} = \frac{k \cdot (k-1) \cdot \dots \cdot (k-m+1)}{m \cdot (m-1) \cdot \dots \cdot 1}$$

$$< \frac{n \cdot (n-1) \cdot \dots \cdot (n-m+1)}{m \cdot (m-1) \cdot \dots \cdot 1} = \binom{n}{m}. \quad \square$$

Example 3.30. Prove that if n, k are positive integers, with k odd, then

$$1 + 2 + \dots + n \mid 1^k + 2^k + \dots + n^k.$$

Proof. This comes down to $n(n+1) \mid 2(1^k + 2^k + \dots + n^k)$. Since $\gcd(n, n+1) = 1$, it suffices to prove separately that $n \mid 2(1^k + \dots + n^k)$ and $n+1 \mid 2(1^k + \dots + n^k)$. But

$$\begin{aligned} 2(1^k + \dots + n^k) &= (1^k + (n-1)^k) + (2^k + (n-2)^k) + \dots + ((n-1)^k + 1^k) + 2n^k \\ &= (1^k + n^k) + (2^k + (n-1)^k) + \dots + (n^k + 1^k) \end{aligned}$$

and we conclude in both cases using the fact that $a + b \mid a^k + b^k$ when k is odd. \square

Example 3.31. (IMC 2012) Is the set of positive integers n for which $n! + 1$ divides $(2012n)!$ finite or infinite?

Proof. The solution is very short but very tricky: we will prove that the set is finite. Write for simplicity $2012 = k$. Assume that $n! + 1 \mid (kn)!$ for some n . Since $n!^k \mid (kn)!$ (this follows by repeated applications of the divisibility $a!b! \mid (a+b)!$) and since $n! + 1$ and $n!^k$ are relatively prime, we must have

$$f(n) = \frac{(kn)!}{n!^k(n! + 1)} \in \mathbf{Z}.$$

However,

$$\begin{aligned} \frac{f(n+1)}{f(n)} &= \frac{(kn+1)(kn+2)\dots(kn+k)(n!+1)}{(n+1)^k((n+1)!+1)} \\ &< \frac{(kn+k)^k}{(n+1)^k} \cdot \frac{n!+1}{(n+1)!+1} < \frac{k^k}{n}, \end{aligned}$$

since

$$\frac{n!+1}{(n+1)!+1} < \frac{1}{n},$$

this last inequality being equivalent to $n! > n - 1$. Thus, if $n > k^k$, then $f(n+1) < f(n)$. Now, if the problem had infinitely many solutions, there would be an infinite decreasing sequence of positive integers, which is clearly absurd. Hence the set is finite. \square

Yet another result that is very useful in practice and follows directly from the previous ones is:

Theorem 3.32. *Let n be an integer greater than 1 and let a be an integer. Then $0, a, 2a, \dots, (n-1)a$ is a complete residue system modulo n if and only if $\gcd(a, n) = 1$.*

Proof. Suppose that $\gcd(a, n) = 1$. It suffices to show that the remainders of $0, a, 2a, \dots, (n-1)a$ when divided by n are pairwise distinct, as this implies that they must be a permutation of $0, 1, 2, \dots, n-1$. If ia and ja give the same remainder when divided by n , then $n \mid (i-j)a$ and by Gauss' lemma we have $n \mid i-j$, which is impossible if $0 \leq i \neq j < n$.

Suppose now that $0, a, 2a, \dots, (n-1)a$ is a complete residue system modulo n , in particular there is $j \in \{1, 2, \dots, n-1\}$ such that $ja \equiv 1 \pmod{n}$, hence $\gcd(a, n) = 1$. \square

Here are two illustrations of theorem 3.32.

Example 3.33. (Gauss) Let a, b be relatively prime integers greater than 1. Prove that

$$\sum_{k=1}^{b-1} \left\lfloor \frac{ka}{b} \right\rfloor = \frac{(a-1)(b-1)}{2}.$$

Proof. Writing $ka = q_k b + r_k$ with $0 \leq r_k < b$, we know that r_1, \dots, r_{b-1} is a permutation of $1, \dots, b-1$, since $\gcd(a, b) = 1$. Thus

$$\sum_{k=1}^{b-1} ka = b \cdot \sum_{k=1}^{b-1} q_k + \sum_{k=1}^{b-1} r_k$$

and so

$$\sum_{k=1}^{b-1} q_k = \frac{1}{b} \sum_{k=1}^{b-1} (ak - k) = \frac{a-1}{b} \cdot \frac{b(b-1)}{2} = \frac{(a-1)(b-1)}{2}.$$

Since $q_k = \left\lfloor \frac{ka}{b} \right\rfloor$, the result follows. \square

Example 3.34. (Landau's identity) Prove that if $m, n > 1$ are relatively prime odd integers, then

$$\sum_{k=1}^{\frac{m-1}{2}} \left\lfloor \frac{kn}{m} \right\rfloor + \sum_{k=1}^{\frac{n-1}{2}} \left\lfloor \frac{km}{n} \right\rfloor = \frac{(m-1)(n-1)}{4}.$$

Proof. Consider the set A of numbers of the form $xm - yn$ with $1 \leq x \leq \frac{n-1}{2}$ and $1 \leq y \leq \frac{m-1}{2}$. We will count the number of elements of this set in two different ways. First, we claim that A has $\frac{(m-1)(n-1)}{4}$ elements, which reduces to checking that the previous numbers are pairwise distinct. But if $xm - yn = x_1m - y_1n$ then $(x - x_1)m = (y - y_1)n$ and so $n \mid m(x - x_1)$. Since $\gcd(m, n) = 1$, we deduce that $n \mid x - x_1$ and since $1 \leq x, x_1 \leq \frac{n-1}{2}$, we must have $x = x_1$ and $y = y_1$, proving the claim.

On the other hand, let us see how many nonnegative numbers are in A . The inequality $xm \geq yn$ is equivalent to $y \leq \frac{xm}{n}$ or $y \leq \left\lfloor \frac{xm}{n} \right\rfloor$. For a given $x \in \{1, 2, \dots, \frac{n-1}{2}\}$ we have $\left\lfloor \frac{xm}{n} \right\rfloor \leq \frac{m-1}{2}$, thus the number of $y \in \{1, 2, \dots, \frac{m-1}{2}\}$ such that $y \leq \frac{xm}{n}$ is $\left\lfloor \frac{xm}{n} \right\rfloor$. Summing over all values of x , we deduce that there are $\sum_{x=1}^{\frac{n-1}{2}} \left\lfloor \frac{xm}{n} \right\rfloor$ nonnegative numbers in A . A similar argument shows that there are $\sum_{y=1}^{\frac{m-1}{2}} \left\lfloor \frac{yn}{m} \right\rfloor$ nonpositive elements in A . We would have a problem if 0 was in A , since it would be counted twice. However, $0 \notin A$, since if $xm = yn$ then $m \mid yn$, then $m \mid y$ and this is impossible since $1 \leq y < m$. Thus $0 \notin A$ and so A has $\sum_{k=1}^{\frac{m-1}{2}} \left\lfloor \frac{kn}{m} \right\rfloor + \sum_{k=1}^{\frac{n-1}{2}} \left\lfloor \frac{km}{n} \right\rfloor$ elements. Combining this with the first paragraph yields the desired result. \square

We end this section with another very useful result, that will be constantly used when dealing with expressions of the form $a^n - b^n$. It is a simple combination of Bézout's theorem and Gauss' lemma, but it is remarkably efficient in practice.

Proposition 3.35. Let a, b and m, n be positive integers. If $\gcd(a, b) = 1$, then

$$\gcd(a^m - b^m, a^n - b^n) = a^{\gcd(m, n)} - b^{\gcd(m, n)}.$$

Proof. Replacing a, b, m, n by $a^{\gcd(m,n)}, b^{\gcd(m,n)}, \frac{m}{\gcd(m,n)}$ and $\frac{n}{\gcd(m,n)}$ respectively, we may assume that $\gcd(m, n) = 1$. Since $a \equiv b \pmod{a-b}$, we have $a^k \equiv b^k \pmod{a-b}$ for all $k \geq 1$. Hence $a-b$ divides $\gcd(a^m - b^m, a^n - b^n)$. Conversely, let $d = \gcd(a^m - b^m, a^n - b^n)$ and let us prove that $d \mid a-b$. We have $a^m \equiv b^m \pmod{d}$ and $a^n \equiv b^n \pmod{d}$, hence $a^{mk} \equiv b^{mk} \pmod{d}$ and $a^{nl} \equiv b^{nl} \pmod{d}$ for all $k, l \geq 1$. Since $\gcd(m, n) = 1$, Bézout's lemma (more precisely corollary 3.19) gives us $k, l \geq 1$ such that $km = ln + 1$. Hence

$$a^{ln+1} = a^{mk} \equiv b^{mk} = b^{nl+1} \equiv b \cdot a^{nl} \pmod{d},$$

that is $d \mid a^{nl}(a-b)$. But since $\gcd(a, b) = 1$, we have $\gcd(a, b^m) = 1$ and hence $\gcd(a, a^m - b^m) = 1$. Since d divides $a^m - b^m$, we conclude that $\gcd(a, d) = 1$. Thus using Gauss' lemma, we obtain $d \mid a-b$ and the result follows. \square

Corollary 3.36. *Let $a > b > 0$ and m, n be positive integers. If $\gcd(a, b) = 1$, then $a^m - b^m$ divides $a^n - b^n$ if and only if $m \mid n$.*

Note that one implication of this corollary is a direct consequence of the fact that if $n = md$, then $a^m - b^m$ divides $a^n - b^n = (a^m)^d - (b^m)^d$. The other implication can also be proved independently of the previous proposition (whose proof is rather technical but whose result is stronger). Indeed, suppose that $a^m - b^m \mid a^n - b^n$ and write $n = mq + r$ for some integers q, r with $0 \leq r < m$. Suppose that $r > 0$. Then

$$a^n - b^n = a^{mq}(a^r - b^r) + b^r(a^{mq} - b^{mq}).$$

By the first step $a^m - b^m \mid a^{mq} - b^{mq}$, hence $a^m - b^m \mid a^{mq}(a^r - b^r)$. Since $\gcd(a, b) = 1$, we have $\gcd(a^{mq}, a^m - b^m) = 1$ and using Gauss' lemma we obtain $a^n - b^n \mid a^r - b^r$. But this is impossible, since $0 < a^r - b^r < a^m - b^m$ (to see why the inequality $a^r - b^r < a^m - b^m$ holds, write it in the form $a^{r-1} + a^{r-2}b + \dots + b^{r-1} < a^{m-1} + \dots + b^{m-1}$).

Let's give a few examples of applications of the previous proposition and corollary:

Example 3.37. Let n be an integer greater than 1. Find all positive integers m such that $(2^n - 1)^2 \mid 2^m - 1$.

Proof. Let m be a solution of the problem. We have $2^n - 1 \mid 2^m - 1$, thus $n \mid m$. Write $m = kn$ for some positive integer k . Then

$$(2^n - 1)^2 \mid 2^{kn} - 1 = (2^n - 1)(1 + 2^n + \dots + (2^n)^{k-1})$$

and so

$$2^n - 1 \mid 1 + 2^n + \dots + (2^n)^{k-1}.$$

On the other hand,

$$1 + 2^n + \dots + (2^n)^{k-1} \equiv 1 + 1 + \dots + 1 = k \pmod{2^n - 1},$$

hence we must have $2^n - 1 \mid k$ and thus $n(2^n - 1) \mid m$. Conversely, if $m = kn$ with $2^n - 1 \mid k$, then the previous computations and congruences show that $(2^n - 1)^2 \mid 2^m - 1$. Thus the solutions of the problem are all multiples of $n(2^n - 1)$. \square

Example 3.38. (Kvant M 1858) Let a and b be positive integers such that

$$\gcd(2a + 1, 2b + 1) = 1.$$

Find the possible values of $\gcd(2^{2a+1} + 2^{a+1} + 1, 2^{2b+1} + 2^{b+1} + 1)$.

Proof. The key observation is that for all $k \geq 0$

$$(2^{2k+1} + 2^{k+1} + 1)(2^{2k+1} - 2^{k+1} + 1) = (2^{2k+1} + 1)^2 - (2^{k+1})^2 = 2^{4k+2} + 1.$$

Set $d = \gcd(2^{2a+1} + 2^{a+1} + 1, 2^{2b+1} + 2^{b+1} + 1)$. Then d divides $2^{4a+2} + 1$ and hence also $2^{8a+4} - 1$. Analogously d divides $2^{8b+4} - 1$. Using the hypothesis, we obtain

$$\gcd(2^{8a+4} - 1, 2^{8b+4} - 1) = 2^{\gcd(8a+4, 8b+4)} - 1 = 2^4 - 1 = 15$$

and d is a divisor of 15. Note that 3 does not divide d since $2^{2a+1} + 2^{a+1} + 1 \equiv 2^{a+1} \pmod{3}$ is not divisible by 3. Thus either $d = 1$ or $d = 5$ and both cases are possible. Indeed, to achieve $\gcd(2^{2a+1} + 2^{a+1} + 1, 2^{2b+1} + 2^{b+1} + 1) = 1$ simply take $a = 1$ and $b = 2$, and to achieve $\gcd(2^{2a+1} + 2^{a+1} + 1, 2^{2b+1} + 2^{b+1} + 1) = 5$ take $a = 3$ and $b = 4$. \square

3.2 Applications to diophantine equations and approximations

The goal of this section is to illustrate the power of the techniques and results established in the previous section, by applying them to the resolution of certain classical diophantine equations. Along the way we will discuss the important topic of approximations of real numbers with rational numbers and its arithmetic applications.

3.2.1 Linear diophantine equations

The simplest diophantine equations are the linear ones. These are the equations of the form

$$a_1x_1 + \dots + a_nx_n = b,$$

where a_1, \dots, a_n, b are given integers. For these equations we have a complete theory, which describes when these equations have solutions as well as methods of finding all solutions.

Theorem 3.39. *Let a_1, \dots, a_n, b be integers. The equation*

$$a_1x_1 + \dots + a_nx_n = b$$

has integral solutions if and only if $\gcd(a_1, \dots, a_n) \mid b$.

Proof. Let $d = \gcd(a_1, \dots, a_n)$. If d does not divide b , then clearly the equation has no integral solution. Assume that $d \mid b$. By Bézout's theorem there are integers y_1, \dots, y_n such that

$$d = a_1y_1 + \dots + a_ny_n.$$

But then setting $x_i = \frac{b}{d} \cdot y_i$ yields an integral solution of the equation. \square

How can we find all solutions of the previous equation? By induction on n we are reduced to discussing the case $n = 2$, which is dealt with in the next theorem.

Theorem 3.40. Let a, b, c be integers with $(a, b) \neq (0, 0)$. Suppose that the equation $ax + by = c$ has integral solutions (which is equivalent to $\gcd(a, b) \mid c$ by the previous theorem) and let (x_0, y_0) be a solution. Then the solutions of the equation are given by

$$\left(x_0 + \frac{b}{\gcd(a, b)}t, y_0 - \frac{a}{\gcd(a, b)}t \right),$$

with $t \in \mathbb{Z}$.

Proof. One easily checks that $\left(x_0 + \frac{b}{\gcd(a, b)}t, y_0 - \frac{a}{\gcd(a, b)}t \right)$ is a solution of the equation for all integers t . Assume now that (x, y) is a solution of the equation. Subtracting the relations $ax + by = c$ and $ax_0 + by_0 = c$ yields $a(x - x_0) = b(y_0 - y)$. Writing $a = du$ and $b = dv$, where $d = \gcd(a, b)$ and $\gcd(u, v) = 1$, we obtain $u(x - x_0) = v(y_0 - y)$. Since $u \mid v(y_0 - y)$ and $\gcd(u, v) = 1$, by Gauss' lemma we can find an integer t such that $y_0 - y = ut$. Then $x - x_0 = vt$, hence $x = x_0 + vt$ and $y = y_0 - ut$. The proof is therefore finished. \square

Example 3.41. Solve in integers the linear diophantine equations

a) $15x + 84y = 39$.

b) $3x + 4y + 5z = 6$.

Proof. a) The equation is equivalent to $5x + 28y = 13$. A solution is $y = 1$, $x = -3$. All solutions are of the form $x = -3 + 28t$, $y = 1 - 5t$, $t \in \mathbb{Z}$, by theorem 3.40.

b) The equation can be written as $3x + 4y = 6 - 5z$. Since $\gcd(3, 4) = 1$ solutions exist for all z , hence we can set $z = s$ for any $s \in \mathbb{Z}$. A solution of $3x + 4y = 1$ is $x = -1$, $y = 1$. So a solution of $3x + 4y = 6 - 5s$ is $x_0 = 5s - 6$, $y_0 = 6 - 5s$. Hence (using again theorem 3.40) all solutions are

$$\begin{cases} x = 5s - 6 + 4t \\ y = 6 - 5s - 3t \\ z = s \end{cases}$$

\square

Example 3.42. (Sylvester 1884) Let $a, b > 1$ be relatively prime integers. Then $ab - a - b$ is the largest integer that cannot be written as $ax + by$, with x, y nonnegative integers.

Proof. Suppose that $ab - a - b = ax + by$ for some nonnegative integers x, y . Then $-b \equiv by \pmod{a}$ and since $\gcd(a, b) = 1$, we have $y \equiv -1 \pmod{a}$. Similarly $x \equiv -1 \pmod{b}$. We deduce that $x \geq b - 1$ and $y \geq a - 1$, hence

$$ab - a - b = ax + by \geq a(b - 1) + b(a - 1) = 2ab - a - b,$$

clearly impossible.

It remains to prove that any integer $n > ab - a - b$ can be written in the desired form. Since $\gcd(a, b) = 1$, there are integers u, v such that $au + bv = n$. Moreover, by replacing u by $u + bt$ and v by $v - at$ for some integer t , we may suppose that $0 \leq u < b$. Then

$$ab - a - b + 1 \leq n = au + bv \leq a(b - 1) + bv,$$

hence $v \geq 0$. The result follows. \square

Example 3.43. Let a_1, \dots, a_n be positive integers and let $\gcd(a_1, \dots, a_n) = k$. Then all sufficiently large multiples N of k can be written $a_1x_1 + \dots + a_nx_n$ with x_1, \dots, x_n positive integers.

Proof. We will prove the statement by induction, the case $n = 1$ being clear. Assume that the result holds for $n - 1$ and let us prove it for n . Fix $a_1, \dots, a_n > 0$ and let $k = \gcd(a_1, \dots, a_n)$ and $l = \gcd(a_1, \dots, a_{n-1})$. Then $k = \gcd(l, a_n)$ by theorem 3.6. Let $N > la_n$ be a multiple of k . Theorem 3.2 shows the existence of an integer x_n such that $N \equiv x_na_n \pmod{l}$. Adding a large multiple of l to x_n , we may assume that $x_n > 0$. Choose the smallest such $x_n > 0$ and observe that $x_n \leq l$ since if $x_n > l$ then $x_n - l$ is a smaller positive solution of the previous congruence. Choose M such that any multiple of l greater than M can be written $x_1a_1 + \dots + x_{n-1}a_{n-1}$ with positive integers x_1, \dots, x_{n-1} (this is possible by the inductive hypothesis). Then for any $N > M + a_nl$ which is a multiple of k , $N - a_nx_n$ is a multiple of l greater than or equal to $N - a_nl > M$, thus we can write $N - a_nx_n = x_1a_1 + \dots + x_{n-1}a_{n-1}$ and so $N = a_1x_1 + \dots + a_nx_n$ with $x_1, \dots, x_n > 0$. This finishes the inductive step and solves the problem. \square

Remark 3.44. If a_1, \dots, a_n are relatively prime positive integers, let $g(a_1, \dots, a_n)$ be the greatest positive integer N for which the equation

$$a_1x_1 + \dots + a_nx_n = N$$

has no solutions in nonnegative integers. Then $g(a_1, \dots, a_n)$ is well-defined by example 3.43. The problem of determining $g(a_1, \dots, a_n)$ is known as the Frobenius coin problem and it is still open except for $n = 2$ (in which case example 3.42 shows that $g(a_1, a_2) = a_1a_2 - a_1 - a_2$).

Example 3.45. (Iran 2002) Let S be a set of positive integers such that $a+b \in S$ whenever $a, b \in S$. Prove that there are positive integers k and N such that for all $n > N$ we have $n \in S$ if and only if $k \mid n$.

Proof. It is clear that S is infinite. Let $a_1 < a_2 < \dots$ be the elements of S and consider the sequence $g_n = \gcd(a_1, \dots, a_n)$. Clearly $g_n \geq g_{n+1}$ for all n , thus the sequence $(g_n)_{n \geq 1}$ is eventually constant, say with value k . Clearly k divides all elements of S . It suffices therefore to prove that all sufficiently large multiples of k are in S . Since S is stable under addition, S contains $a_1x_1 + \dots + a_nx_n$ for any $a_1, \dots, a_n \in S$ and x_1, \dots, x_n positive integers. The result follows then from example 3.43. \square

3.2.2 Pythagorean triples

We want to discuss now one of the most classical and important diophantine equations, namely

$$x^2 + y^2 = z^2.$$

Triples of integers (x, y, z) satisfying this equation are called Pythagorean. Finding Pythagorean triples is equivalent to finding right-angled triangles with integer side-lengths. In order to describe all solutions of this equation, we will need the following result, which turns out to be extremely useful in the study of diophantine equations.

Theorem 3.46. *Let a, b be relatively prime positive integers such that $ab = c^n$ for some positive integer c . Then a and b are both n th powers of positive integers.*

Proof. Let $d = \gcd(a, c)$ and write $a = du$ and $c = dv$ for relatively prime positive integers u, v . Then $ub = d^{n-1}v^n$. Since $\gcd(u, v) = 1$, we have $\gcd(u, v^n) = 1$ (corollary 3.26). Since $u \mid d^{n-1}v^n$, Gauss' lemma yields $u \mid d^{n-1}$ and so $v^n = \frac{u}{d^{n-1}}b$ is a multiple of b . On the other hand $v^n \mid ub$ and $\gcd(v^n, u) = 1$, thus using again Gauss' lemma we obtain $v^n \mid b$. We conclude that $b = v^n$ and $u = d^{n-1}$, thus $a = d^n$. The result follows. \square

Before dealing with the resolution of the equation $x^2 + y^2 = z^2$ we would like to illustrate the previous theorem with a few interesting examples.

Example 3.47. Prove that the product of three consecutive positive integers is never a perfect power.

Proof. Write the three consecutive integers $n-1, n, n+1$, and suppose that $(n-1)n(n+1) = a^d$, with $a, d > 1$.

Then $n(n^2-1) = a^d$ and since $\gcd(n, n^2-1) = 1$, it follows that both n and n^2-1 are d th powers. Say $n = c^d$ and $n^2-1 = e^d$, for some integers $c, e > 1$. Then $c^{2d} - e^d = 1$, which can also be written as $(c^2 - e)(c^{2(d-1)} + \dots + e^{d-1}) = 1$. This is clearly impossible, since $c^2 - e \geq 1$ and $c^{2(d-1)} + \dots + e^{d-1} \geq d > 1$. \square

Example 3.48. (IMO Shortlist 2007) Let b, n be integers greater than 1 such that for all $k > 1$ one can find an integer a such that $k \mid b - a^n$. Prove that b is the n th power of an integer.

Proof. Choosing $k = b^2$, it follows that there are integers a and c such that $b - a^n = cb^2$. This can be written as $b(1 - cb) = a^n$. Thus b and $1 - cb$ are positive numbers, relatively prime and whose product is an n th power. It follows that both are n th powers. In particular, b is an n th power, as desired. \square

Example 3.49. (Vietnam 2013) Find all integers x such that $\frac{x^{1000}-1}{x-1}$ is a perfect square.

Proof. Clearly $x = -1$ and $x = 0$ are solutions, and we will prove that they are the only solutions of the problem. If $x < -1$, then the fraction is negative

and cannot be a perfect square, so we need only consider $x > 1$. Since

$$\frac{x^{1000} - 1}{x - 1} = \frac{x^{500} - 1}{x - 1} \cdot (x^{500} + 1)$$

and $\gcd\left(\frac{x^{500}-1}{x-1}, x^{500} + 1\right) \mid 2$, and since $x^{500} + 1$ is not a square¹, we deduce that $x^{500} + 1 = 2u^2$ and $\frac{x^{500}-1}{x-1} = 2v^2$ for some integers $u, v > 1$. Thus

$$\frac{x^{250} - 1}{x - 1} \cdot \frac{x^{250} + 1}{2} = v^2.$$

Note that 4 does not divide $x^{250} + 1$ (since 4 does not divide $u^2 + 1$ for any integer u), hence $\frac{x^{250}-1}{x-1}$ and $\frac{x^{250}+1}{2}$ are relatively prime and so each of them must be a square. But then $\frac{x^{125}-1}{x-1} \cdot (x^{125} + 1)$ is a square and $x^{125} + 1$ and $\frac{x^{125}-1}{x-1}$ are relatively prime (since $\frac{x^{125}-1}{x-1}$ is odd, because x is odd). Thus $x^{125} + 1$ is a square, say $x^{125} + 1 = z^2$. Then $(z - 1)(z + 1) = x^{125}$ and $z - 1, z + 1$ are relatively prime (since z is even), thus $z - 1$ and $z + 1$ are both the 125th power of some positive integers p, q . But then $q^{125} - p^{125} = 2$ and so $q - p = 1$ or $q - p = 2$. In both case it is easy to see that we cannot have $q^{125} - p^{125} = 2$. Thus the only solution is $x = 0$. \square

We are now ready to describe all Pythagorean triples.

Theorem 3.50. *The solutions in positive integers of the equation*

$$x^2 + y^2 = z^2$$

are given by

$$x = d(m^2 - n^2), \quad y = 2dmn, \quad z = (m^2 + n^2)d$$

or

$$x = 2dmn, \quad y = (m^2 - n^2)d, \quad z = (m^2 + n^2)d,$$

where $m > n > 0$ are relatively prime and of different parity, and where $d > 0$.

¹Since $x^{500} + 1$ lies strictly between the consecutive squares x^{500} and $(x^{250} + 1)^2$.

Proof. It is not difficult to check that the given triples are solutions of the equation: this reduces to the equality

$$(m^2 - n^2)^2 + (2mn)^2 = (m^2 + n^2)^2,$$

which is easy to check.

Conversely, let (x, y, z) be a solution of the equation with $x, y, z > 0$ and let $d = \gcd(x, y)$, so that $x = da$ and $y = db$ with a, b relatively prime positive integers. Moreover,

$$d^2(a^2 + b^2) = z^2$$

hence $d^2 \mid z^2$ and so $d \mid z$. Say $z = dc$ for some positive integer c , then

$$a^2 + b^2 = c^2.$$

Since a, b are relatively prime, the previous relation implies that a, b, c are pairwise relatively prime. Also, note that c is odd: otherwise, since a, b are relatively prime they must be both odd but then $c^2 = a^2 + b^2 \equiv 2 \pmod{4}$, impossible. Thus a and b have different parities. By symmetry, we may assume that a is odd and b is even. Rewrite the equality $a^2 + b^2 = c^2$ as

$$\left(\frac{b}{2}\right)^2 = \frac{c-a}{2} \frac{c+a}{2}$$

and observe that since $\gcd(a, c) = 1$ we also have $\gcd\left(\frac{c-a}{2}, \frac{c+a}{2}\right) = 1$ (note that the sum and difference of the numbers $\frac{c+a}{2}$ and $\frac{c-a}{2}$ are c and a respectively). We deduce that $\frac{c-a}{2}$ and $\frac{c+a}{2}$ are both perfect squares, say

$$\frac{c-a}{2} = n^2, \quad \frac{c+a}{2} = m^2.$$

Note that $m > n$ are relatively prime positive integers of different parities (since $m^2 + n^2 = c$ is odd). Also $b = 2mn$ and

$$x = d(m^2 - n^2), \quad y = 2dmn, \quad z = d(m^2 + n^2),$$

as desired. □

One of the most famous problems in number theory is the resolution of Fermat's equation

$$x^n + y^n = z^n.$$

We have just seen how to solve it for $n = 2$. The general case was solved by Wiles in 1994 (more than 350 years after the problem was posed), who proved that for $n > 2$ there are no nontrivial solutions. The proof of this deep result is one of the most spectacular applications of the interplay between number theory and algebraic geometry (needless to say, the proof goes far beyond the scope of this modest book). Already the case $n = 3$ is quite challenging (even though in this case there is an elementary, though fairly technical proof). The next theorem deals with the case $n = 4$ and establishes a stronger result, using Fermat's method of infinite descent (we have already encountered some applications of this method in the first chapter).

Theorem 3.51. (*Fermat*) *The equations $x^4 + y^4 = z^2$ and $x^4 - y^4 = z^2$ have no nontrivial (i.e. with $xyz \neq 0$) integral solutions.*

Proof. We only give the proof for the equation $x^4 + y^4 = z^2$, the argument being similar for the other one. We may restrict ourselves to solutions x, y, z in which $x, y, z \geq 0$ (since changing each of x, y, z into its absolute value does not change the fact that they form a nontrivial solution of the equation). Assume the contrary and consider a nontrivial solution (x_0, y_0, z_0) with smallest possible value of z_0 . Then necessarily $\gcd(x_0, y_0) = 1$ (otherwise letting $d = \gcd(x_0, y_0)$, d^2 must divide z and so $(\frac{x_0}{d}, \frac{y_0}{d}, \frac{z_0}{d^2})$ gives a nontrivial solution with smaller value of z , contradicting the minimality of z_0). Also, one of x_0, y_0 must be even (otherwise we obtain $z^2 \equiv 2 \pmod{4}$, a contradiction), say without loss of generality y_0 is even. Using theorem 3.50 we may find relatively prime positive integers a, b such that

$$x_0^2 = a^2 - b^2, y_0^2 = 2ab, z_0 = a^2 + b^2.$$

Since $x_0^2 = a^2 - b^2$, x_0 is odd (as y_0 is even and $\gcd(x_0, y_0) = 1$) and $\gcd(a, b) = 1$, we deduce, using again theorem 3.50, the existence of relatively prime positive integers c, d such that

$$x_0 = c^2 - d^2, b = 2cd, a = c^2 + d^2.$$

It follows that

$$cd(c^2 + d^2) = \frac{ab}{2} = \left(\frac{y_0}{2}\right)^2.$$

Since $c, d, c^2 + d^2$ are pairwise relatively prime, we conclude that each of them must be a perfect square, say $c = u^2, d = v^2$ and $c^2 + d^2 = w^2$. Then

$$u^4 + v^4 = w^2$$

and so (u, v, w) is a nontrivial solution. By minimality of z_0 , this forces $w \geq z_0$. But this is certainly impossible, since

$$z_0 = a^2 + b^2 > a^2 = (c^2 + d^2)^2 > c^2 + d^2 = u^4 + v^4 = w^2.$$

The result follows. □

Remark 3.52. 1. On the other hand, the equation

$$x^4 + y^4 + z^4 = t^4$$

has nontrivial solutions: a famous example due to Elkies (1988) is

$$2682440^4 + 15365639^4 + 18796760^4 = 20615673^4.$$

Another example, found by Frye is

$$95800^4 + 217519^4 + 414560^4 = 422481^4.$$

The equation

$$x^5 + y^5 + z^5 + t^5 = w^5$$

also has nontrivial solutions, for instance Lander and Parkin (1967) found the solution

$$144^5 = 27^5 + 84^5 + 110^5 + 133^5.$$

These examples disprove a conjecture of Euler, namely that for $n > 2$ the equation

$$a_1^n + a_2^n + \dots + a_{n-1}^n = b^n$$

has no solutions in positive integers (this turns out to be true for $n = 3$, as we have already mentioned).

2. With exactly the same arguments one can prove that the equation $x^4 - y^4 = z^2$ has no nontrivial integral solutions. We deduce immediately that the equation $x^4 + y^4 = 2z^2$ has only the obvious integral solutions, by writing it in the form

$$z^4 - (xy)^4 = \left(\frac{x^4 - y^4}{2} \right)^2.$$

3. In general, if d is a positive integer, then one can prove that the equation $x^4 - y^4 = dz^2$ either has no nontrivial solutions or infinitely many solutions in relatively prime positive integers.

Example 3.53. For which integers x, y do we have $x^4 - 2y^2 = 1$?

Proof. Writing the equation as

$$x^4 + y^4 = (y^2 + 1)^2$$

and applying Fermat's theorem above we obtain $y = 0$ and then $x = \pm 1$. \square

Example 3.54. Find all integers x, y such that $8x^4 + 1 = y^2$.

Proof. Suppose that (x, y) is a solution. Replacing x, y with their absolute values, we may assume that $x, y \geq 0$. If $y = 1$, we obtain the solution $(0, 1)$. Suppose that $y > 1$. Clearly y is odd, say $y = 2z + 1$ for some positive integer z . Then $z(z + 1) = 2x^4$ and since $\gcd(z, z + 1) = 1$, we deduce that either $z = 2a^4$ and $z + 1 = b^4$ for some positive integers a, b , or $z = a^4$ and $z + 1 = 2b^4$ for some positive integers a, b . In the first case we obtain $b^4 - 2a^4 = 1$, contradicting the result established in the previous example. In the second case we obtain $a^4 + 1 = 2b^4$, which can be written as

$$a^4 + \left(\frac{a^4 - 1}{2} \right)^2 = b^8.$$

Since $a, b \geq 1$ and the equation $x^4 + y^2 = z^4$ is impossible in positive integers, Since the equation $x^4 + z^2 = y^4$ has only trivial solutions and since we deduce that $a = 1$ and $b = 1$, thus $z = 1$ and $x = 1, y = 3$. We conclude that the only solutions are $(x, y) = (0, \pm 1), (\pm 1, \pm 3)$. \square

Example 3.55. Solve in integers the equation

$$x^4 + (x^2 + 1)^2 = y^2.$$

Proof. Again, we may assume that $x, y \geq 0$. If $x = 0$, we obtain $y = 1$, so assume that $x > 0$. Then $x^2, x^2 + 1$ and y form a Pythagorean triple with $\gcd(x^2, x^2 + 1) = 1$ and moreover x^2 is even (if x is odd, then the left-hand side is congruent to 5 mod 8, while the right-hand side is congruent to 1 mod 8). Thus there are relatively prime positive integers $m > n$ of different parity such that $x^2 = 2mn$ and $x^2 + 1 = m^2 - n^2$. Letting $x = 2a$, we obtain $mn = 2a^2$ and $m^2 - n^2 = 4a^2 + 1$. Since m, n have different parity and $m^2 - n^2 \equiv 1 \pmod{4}$, m must be odd and n must be even. Since $mn = 2a^2$ and $\gcd(m, n) = 1$, we conclude that $n = 2u^2$ and $m = v^2$ for some integers $u, v > 0$, and $a = uv$. We conclude that

$$v^4 - 4u^4 = 4u^2v^2 + 1,$$

which can also be written as

$$(v^2 - 2u^2)^2 - 8u^4 = 1.$$

By the previous example we obtain (since $u > 0$) $u = 1$ and $v^2 - 2u^2 = \pm 3$. This is however impossible, thus the only solution is $(0, \pm 1)$. \square

Example 3.56. Solve in integers the equation

$$(2x^2 - 1)^2 = 2y^2 - 1.$$

Proof. We may assume that $x, y \geq 0$. Clearly $y \geq 1$ and if $y = 1$ we obtain $x = 0$ or $x = 1$. Assume from now on that $y > 1$, so that $x > 1$.

Write the equation as

$$(x^2)^2 + (x^2 - 1)^2 = y^2.$$

We discuss two cases, according to the parity of x .

Suppose that x is odd, then $x^2 = a^2 - b^2$, $x^2 - 1 = 2ab$ and $y = a^2 + b^2$ for some $a > b > 0$ relatively prime and of different parity. Write $a - b = u^2$ and $a + b = v^2$ with $0 < u < v$ relatively prime and odd (note that such u, v

exist since $(a-b)(a+b) = x^2$ and $\gcd(a-b, a+b) = 1$. Then $x = uv$ and the equation $x^2 - 1 = 2ab$ becomes

$$(uv)^2 - 1 = 2 \cdot \frac{u^2 + v^2}{2} \cdot \frac{v^2 - u^2}{2}$$

or

$$2u^2v^2 - 2 = v^4 - u^4.$$

This is equivalent to $(v^2 - u^2)^2 = 2(u^4 - 1)$ and writing $v^2 - u^2 = 2w$ yields $u^4 - 2w^2 = 1$. Using example 3.53 we obtain a contradiction.

Suppose now that x is even. Similar arguments yield the existence of $a > b > 0$ relatively prime, of different parity such that $x^2 = 2ab$, $x^2 - 1 = a^2 - b^2$ and $y = a^2 + b^2$. Since $a^2 - b^2 = x^2 - 1 \equiv -1 \pmod{4}$, we deduce that a is even and b is odd. Since $2ab = x^2$ is a square, a is even, b is odd and $\gcd(a, b) = 1$, we obtain $a = 2m^2$, $b = n^2$ and $x = 2mn$ for some positive integers m, n , which are relatively prime. Then the equation $x^2 - 1 = a^2 - b^2$ becomes

$$4m^2n^2 - 1 = 4m^4 - n^4.$$

This can be rewritten as

$$(n^2 + 2m^2)^2 = 8m^4 + 1.$$

Using example 3.54 we obtain $m = 1$ and $n^2 + 2m^2 = 3$, thus $m = n = 1$. But then $a = 2, b = 1, x = 2$ and $y = 5$.

We conclude that the solutions are $(0, \pm 1), (\pm 1, \pm 1), (\pm 2, \pm 5)$. \square

Example 3.57. Find all integers x, y such that

$$1 + x + x^2 + x^3 = y^2.$$

Proof. Write the equation as $(1+x)(1+x^2) = y^2$, which makes it clear that $x \geq -1$. If $x = -1$ we obtain the solution $(-1, 0)$, and if $x = 0$ we obtain the solutions $(0, \pm 1)$. If $x = 1$ we obtain the solutions $(1, \pm 2)$.

Assume from now on that $x > 1$ and, without loss of generality, that $y \geq 0$ (and so $y > 2$). If x is even, then $\gcd(1+x, 1+x^2) = 1$ and we deduce that

$1+x$ and $1+x^2$ are perfect square, which is clearly impossible. Thus x is odd, and then $\gcd(1+x, 1+x^2) = 2$. We deduce that $1+x = 2a^2$ and $1+x^2 = 2b^2$ for some $a, b \geq 1$, and $y = 2ab$. But then

$$(2a^2 - 1)^2 = 2b^2 - 1.$$

Using the previous example we obtain $a = 2$ and $b = 5$.

But then $x = 2a^2 - 1 = 7$ and $y = 20$.

We conclude that the solutions of the problem are

$$(-1, 0), (0, \pm 1), (1, \pm 2), (7, \pm 20). \quad \square$$

Example 3.58. (Bulgaria 1998) Prove that the equation $x^2y^2 = z^2(z^2 - x^2 - y^2)$ has no solutions in positive integers.

Proof. Assume the contrary and let $a = x^2 + y^2$ and $b = 2xy$. Then

$$a^2 - b^2 = (x^2 - y^2)^2$$

and

$$a^2 + b^2 = x^4 + y^4 + 6x^2y^2.$$

On the other hand, since the equation $(z^2)^2 - z^2a - \frac{b^2}{4}$ has integer solutions, we deduce that its discriminant $a^2 + b^2$ is a perfect square. Thus $a^2 - b^2$ and $a^2 + b^2$ are both squares and so $a^4 - b^4 = t^2$ for some integer t . Since $a, b > 0$, we deduce that $a = b$ and so $x = y$. But then $(z^2 - x^2)^2 = 2x^4$, contradicting the fact that $\sqrt{2}$ is irrational (see example 3.62 for a proof of a more general result). \square

Remark 3.59. The proof shows that already the equation $x^2y^2 = z(z^2 - x^2 - y^2)$ has no solutions in positive integers.

3.2.3 The rational root theorem

We will discuss now another application of Gauss' lemma, the rational root theorem. This theorem bounds the denominators of the possible rational zeros of a polynomial with integer coefficients. One important consequence is that any rational root of a monic polynomial with integer coefficients must be an integer.

Theorem 3.60. (*the rational root theorem*) Let $f(X) = a_n X^n + \dots + a_0$ be a polynomial with integer coefficients and $a_n \neq 0$. If $x = \frac{p}{q}$ (with p, q relatively prime integers) is a rational root of f , then $q \mid a_n$.

Proof. Multiplying the equality $f(x) = 0$ by q^n yields

$$a_n p^n + a_{n-1} p^{n-1} q + \dots + a_0 q^n = 0.$$

All terms in the sum except the first one are clearly multiples of q . Thus $q \mid a_n p^n$. On the other hand $\gcd(q, p) = 1$, thus $\gcd(q, p^n) = 1$ and using Gauss' lemma we conclude that $q \mid a_n$, as desired. \square

Corollary 3.61. Let f be a monic polynomial (i.e. the leading coefficient of f is 1) with integer coefficients. Any rational root of f is an integer.

Example 3.62. Let n be a positive integer and let $d > 1$ be an integer. Prove that if $\sqrt[d]{n}$ is a rational number, then it is an integer.

Proof. Let $x = \sqrt[d]{n}$ and observe that x is a rational root of the monic polynomial with integer coefficients $X^d - n$. Thus x must be an integer, by corollary 3.61. \square

In particular, if a, b, c are integers with $a \neq 0$ and if the equation $ax^2 + bx + c = 0$ has a rational solution, then the discriminant $\Delta = b^2 - 4ac$ must be a perfect square. Indeed, $\sqrt{\Delta} = |2ax + b|$ is rational and we conclude it is an integer using the previous example. Here is a nice application of this observation.

Example 3.63. (Kvant M 1740) Let a, b, c be positive integers such that

$$a^2 + b^2 + c^2 = (a - b)^2 + (b - c)^2 + (c - a)^2.$$

Prove that ab, bc, ca and $ab + bc + ca$ are all perfect squares.

Proof. We can rewrite the given relation as

$$a^2 + b^2 + c^2 = 2(ab + bc + ca).$$

$1+x$ and $1+x^2$ are perfect square, which is clearly impossible. Thus x is odd, and then $\gcd(1+x, 1+x^2) = 2$. We deduce that $1+x = 2a^2$ and $1+x^2 = 2b^2$ for some $a, b \geq 1$, and $y = 2ab$. But then

$$(2a^2 - 1)^2 = 2b^2 - 1.$$

Using the previous example we obtain $a = 2$ and $b = 5$.

But then $x = 2a^2 - 1 = 7$ and $y = 20$.

We conclude that the solutions of the problem are

$$(-1, 0), (0, \pm 1), (1, \pm 2), (7, \pm 20). \quad \square$$

Example 3.58. (Bulgaria 1998) Prove that the equation $x^2y^2 = z^2(z^2 - x^2 - y^2)$ has no solutions in positive integers.

Proof. Assume the contrary and let $a = x^2 + y^2$ and $b = 2xy$. Then

$$a^2 - b^2 = (x^2 - y^2)^2$$

and

$$a^2 + b^2 = x^4 + y^4 + 6x^2y^2.$$

On the other hand, since the equation $(z^2)^2 - z^2a - \frac{b^2}{4}$ has integer solutions, we deduce that its discriminant $a^2 + b^2$ is a perfect square. Thus $a^2 - b^2$ and $a^2 + b^2$ are both squares and so $a^4 - b^4 = t^2$ for some integer t . Since $a, b > 0$, we deduce that $a = b$ and so $x = y$. But then $(z^2 - x^2)^2 = 2x^4$, contradicting the fact that $\sqrt{2}$ is irrational (see example 3.62 for a proof of a more general result). \square

Remark 3.59. The proof shows that already the equation $x^2y^2 = z(z^2 - x^2 - y^2)$ has no solutions in positive integers.

3.2.3 The rational root theorem

We will discuss now another application of Gauss' lemma, the rational root theorem. This theorem bounds the denominators of the possible rational zeros of a polynomial with integer coefficients. One important consequence is that any rational root of a monic polynomial with integer coefficients must be an integer.

Theorem 3.60. (*the rational root theorem*) Let $f(X) = a_n X^n + \dots + a_0$ be a polynomial with integer coefficients and $a_n \neq 0$. If $x = \frac{p}{q}$ (with p, q relatively prime integers) is a rational root of f , then $q \mid a_n$.

Proof. Multiplying the equality $f(x) = 0$ by q^n yields

$$a_n p^n + a_{n-1} p^{n-1} q + \dots + a_0 q^n = 0.$$

All terms in the sum except the first one are clearly multiples of q . Thus $q \mid a_n p^n$. On the other hand $\gcd(q, p) = 1$, thus $\gcd(q, p^n) = 1$ and using Gauss' lemma we conclude that $q \mid a_n$, as desired. \square

Corollary 3.61. Let f be a monic polynomial (i.e. the leading coefficient of f is 1) with integer coefficients. Any rational root of f is an integer.

Example 3.62. Let n be a positive integer and let $d > 1$ be an integer. Prove that if $\sqrt[d]{n}$ is a rational number, then it is an integer.

Proof. Let $x = \sqrt[d]{n}$ and observe that x is a rational root of the monic polynomial with integer coefficients $X^d - n$. Thus x must be an integer, by corollary 3.61. \square

In particular, if a, b, c are integers with $a \neq 0$ and if the equation $ax^2 + bx + c = 0$ has a rational solution, then the discriminant $\Delta = b^2 - 4ac$ must be a perfect square. Indeed, $\sqrt{\Delta} = |2ax + b|$ is rational and we conclude it is an integer using the previous example. Here is a nice application of this observation.

Example 3.63. (Kvant M 1740) Let a, b, c be positive integers such that

$$a^2 + b^2 + c^2 = (a - b)^2 + (b - c)^2 + (c - a)^2.$$

Prove that ab, bc, ca and $ab + bc + ca$ are all perfect squares.

Proof. We can rewrite the given relation as

$$a^2 + b^2 + c^2 = 2(ab + bc + ca).$$

Considering this as a quadratic equation in a , the discussion preceding the problem shows that the discriminant $\Delta = 16bc$ must be a perfect square. We conclude that bc is a perfect square and by symmetry we also obtain that ab and ac are perfect squares. Writing $bc = x^2$ for some integer x , we obtain $a = b + c \pm 2x$ and so $b + c = a + \varepsilon \cdot 2x$ with $\varepsilon \in \{-1, 1\}$. But then

$$ab + bc + ca = x^2 + a(b + c) = x^2 + a(a + 2\varepsilon \cdot x) = (a + \varepsilon \cdot x)^2,$$

finishing the proof. \square

The following exercise refines the rational root theorem.

Example 3.64. Let f be a polynomial with integer coefficients and let $x = \frac{p}{q}$ be a rational root of f , with p, q relatively prime integers. Then we can find a polynomial g with integer coefficients such that $f(X) = (qX - p)g(X)$.

Proof. Write $f(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_0$ and let us look for g of the form $g(X) = b_{n-1} X^{n-1} + \dots + b_0$. The equality $f(X) = (qX - p)g(X)$ reduces (after looking at the coefficient of X^j on both sides for all j) to the system of equations

$$-pb_0 = a_0, qb_0 - pb_1 = a_1, \dots, qb_{n-2} - pb_{n-1} = a_{n-1}, qb_{n-1} = a_n.$$

Solving successively we obtain

$$b_0 = -\frac{a_0}{p}, b_1 = -\frac{qa_0 + pa_1}{p^2}, \dots,$$

$$b_{n-1} = -\frac{q^{n-1}a_0 + pq^{n-2}a_1 + \dots + p^{n-1}a_{n-1}}{p^n} = \frac{a_n}{q}.$$

Thus we need to prove that all these expressions are integers. Note that the rational root theorem is precisely the statement that b_{n-1} is an integer. In general, we need to show that p^{k+1} divides $q^k a_0 + pq^{k-1} a_1 + \dots + p^k a_k$. However we know that

$$a_n p^n + a_{n-1} p^{n-1} q + \dots + a_{k+1} p^{k+1} q^{n-k-1} + a_k p^k q^{n-k} + \dots + a_0 q^n = 0.$$

We deduce that p^{k+1} divides $a_k p^k q^{n-k} + \dots + a_0 q^n = q^{n-k} (a_k p^k + \dots + a_0 q^k)$. Since p^{k+1} and q^{n-k} are relatively prime, we deduce that p^{k+1} divides $a_k p^k + \dots + a_0 q^k$, as needed. \square

Here is a nice application of the rational root theorem. Assume that a, b are rational numbers such that $a + b$ and ab are integers. We claim that a and b are actually integers. Indeed, a and b are roots of $(X - a)(X - b) = X^2 - (a + b)X + ab$, which is a monic polynomial with integer coefficients, by assumption. Using the previous corollary, we deduce that a and b must be integers. A similar argument shows that if a, b, c are rational numbers and $a + b + c, ab + bc + ca$ and abc are all integers, then a, b, c are all integers. This kind of result is very useful in many contexts, and can lead to quite surprising results.

Example 3.65. Find all positive integers a, b, c such that $\frac{a}{b} + \frac{b}{c} + \frac{c}{a}$ and $\frac{b}{a} + \frac{c}{b} + \frac{a}{c}$ are both integers.

Proof. Consider the polynomial with roots $\frac{a}{b}, \frac{b}{c}, \frac{c}{a}$

$$f(X) = \left(X - \frac{a}{b}\right) \cdot \left(X - \frac{b}{c}\right) \cdot \left(X - \frac{c}{a}\right).$$

A brutal expansion shows that

$$f(X) = X^3 - \left(\frac{a}{b} + \frac{b}{c} + \frac{c}{a}\right)X^2 + \left(\frac{b}{a} + \frac{c}{b} + \frac{a}{c}\right)X - 1$$

and so (a, b, c) is a solution of the problem if and only if f has integer coefficients. Consider such (a, b, c) and note that f is also monic and has rational roots $\frac{a}{b}, \frac{b}{c}, \frac{c}{a}$. By the rational root theorem, these roots must be integers, thus $\frac{a}{b}, \frac{b}{c}, \frac{c}{a}$ are positive integers. Since their product is 1, they all must be 1 and so $a = b = c$. Conversely, if $a = b = c$ then obviously (a, b, c) is a solution of the problem. \square

Example 3.66. (USAMO 2009) Let s_1, s_2, s_3, \dots be an infinite, nonconstant sequence of rational numbers.

Suppose that t_1, t_2, t_3, \dots is also an infinite, nonconstant sequence of rational numbers with the property that $(s_i - s_j)(t_i - t_j)$ is an integer for all i and j . Prove that there is a nonzero rational number r such that $(s_i - s_j)r$ and $(t_i - t_j)/r$ are integers for all i and j .

Proof. By working with the sequence $0, s_2 - s_1, s_3 - s_1, \dots$ instead of s_1, s_2, \dots , we may assume that $s_1 = 0$. Similarly, we may assume that $t_1 = 0$. In particular $s_i t_i$ is an integer for all i , thanks to the assumption of the problem. But then $(s_i - s_j)(t_i - t_j) - (s_i t_i + s_j t_j)$ must be an integer, i.e. $s_i t_j + s_j t_i$ is an integer for all i, j . Fix i, j and note that $(s_i t_j) \cdot (s_j t_i)$ is an integer, since it equals $(s_i t_i) \cdot (s_j t_j)$. Since the sum and product of the rational numbers $s_i t_j$ and $s_j t_i$ are integers, both these rational numbers must be integers. Thus $s_i t_j$ is an integer for all i, j . By choosing i such that $s_i \neq 0$, we see in particular that there is a nonzero integer N such that $N t_j$ is an integer for all j . Define

$$r = \frac{1}{N} \gcd(N t_1, N t_2, \dots).$$

By Bézout's theorem (note that $\gcd(N t_1, N t_2, \dots)$ is actually the gcd of finitely many of the $N t_j$'s) r is a linear combination with integer coefficients of t_1, t_2, \dots (and only finitely many of the coefficients will be nonzero). Since $s_i t_j$ is an integer for all i, j , we deduce that $r s_i$ is an integer for all i , thus $(s_i - s_j)r$ is an integer for all i, j . Finally, it is clear by construction that

$$\frac{t_i - t_j}{r} = \frac{N t_i - N t_j}{\gcd(N t_1, N t_2, \dots)}$$

is an integer for all i, j . □

3.2.4 Farey fractions and Pell's equation

In this subsection we start the study of two fundamental diophantine equations

$$x^2 + y^2 = n \quad \text{and} \quad x^2 - n y^2 = 1,$$

where n is a given positive integer. Note that the first equation clearly has finitely many solutions since $|x|$ and $|y|$ cannot exceed \sqrt{n} . We will prove that the equation $x^2 + y^2 = n$ has as many solutions as the congruence $z^2 \equiv -1 \pmod{n}$. We will see in later chapters how to find the number of solutions of this congruence (once enough theory is introduced, this will become a straightforward exercise, while the problem of giving a closed formula for the number of solutions of the equation $x^2 + y^2 = n$ is definitely not easy). We will also

prove that if n is not a perfect square, then $x^2 - ny^2 = 1$ (known as Pell's equation) has infinitely many solutions and that if one knows the smallest solution, then one can obtain all other solutions by a simple recipe. In order to prove all these results, we will introduce and study a very beautiful object: Farey sequences.

Let $n > 0$ be an integer. Consider all fractions (in lowest form) whose (positive) denominator does not exceed n , in other words all rational numbers of the form $\frac{a}{b}$ with a, b relatively prime integers and $0 < b \leq n$. Arrange these fractions in increasing order and call the resulting sequence the Farey sequence of order n .

The key property of Farey sequences is then:

Theorem 3.67. *Let $\frac{a}{b}$ and $\frac{a'}{b'}$ be consecutive terms in the Farey sequence of order n . Then*

$$b + b' \geq n + 1 \quad \text{and} \quad ba' - ab' = \pm 1.$$

Proof. We may assume that $\frac{a}{b} < \frac{a'}{b'}$. We will actually identify the fraction $\frac{a'}{b'}$ as follows. Consider two integers x, y such that

$$bx - ay = 1 \quad \text{and} \quad -b < y - n \leq 0.$$

Note that such integers exist: by theorem 3.20 the congruence $ay \equiv -1 \pmod{b}$ has at least one solution y in the set $\{n, n-1, \dots, n-b+1\}$ (which is a complete residue system modulo b). Note that since by definition $b \leq n$ and $y > n-b$, we have $y > 0$.

We prove now that $\frac{a'}{b'} = \frac{x}{y}$. Suppose that this is not the case. Since $\frac{a}{b}$ and $\frac{a'}{b'}$ are consecutive in the Farey sequence of order n and since $\frac{x}{y}$ is also a term of this sequence (as clearly $\gcd(x, y) = 1$ and $0 < y \leq n$) and

$$\frac{x}{y} = \frac{a}{b} + \frac{1}{by} > \frac{a}{b},$$

we deduce that $\frac{x}{y} > \frac{a'}{b'}$, hence

$$\frac{x}{y} - \frac{a'}{b'} = \frac{b'x - a'y}{b'y} \geq \frac{1}{b'y}.$$

A similar argument yields

$$\frac{a'}{b'} - \frac{a}{b} \geq \frac{1}{bb'},$$

thus

$$\frac{1}{by} = \frac{x}{y} - \frac{a}{b} \geq \frac{1}{b'y} + \frac{1}{bb'},$$

which gives $b' \geq y + b > n$, a contradiction with the fact that $\frac{a'}{b'}$ is in the Farey sequence of order n .

Thanks to the previous paragraph we know that $\frac{a'}{b'} = \frac{x}{y}$ and so $b'x = a'y$. We deduce from Gauss' lemma that $b' = y$ and then $a' = x$. Taking into account the choice of x, y , we conclude that

$$a'b - ab' = bx - ay = 1 \quad \text{and} \quad b + b' = b + y > n.$$

The result follows. □

A simple but very important consequence of the previous theorem is the following approximation result:

Corollary 3.68. *If x is a real number and n is a positive integer, then we can find relatively prime integers a, b such that $0 < b \leq n$ and*

$$\left| x - \frac{a}{b} \right| \leq \frac{1}{b(n+1)}.$$

Proof. Let $f_1 < f_2 < \dots < f_d$ be the terms of the Farey sequence of order n that belong to the closed interval $[x - 1, x + 1]$. If $f_i = \frac{a_i}{b_i}$ with $\gcd(a_i, b_i) = 1$ and $0 < b_i \leq n$, consider

$$g_i = \frac{a_i + a_{i+1}}{b_i + b_{i+1}}$$

for $1 \leq i < d$. Thanks to the previous theorem we have

$$g_i - f_i = \frac{b_i a_{i+1} - a_i b_{i+1}}{b_i(b_i + b_{i+1})} = \frac{1}{b_i(b_i + b_{i+1})} \in \left(0, \frac{1}{(n+1)b_i} \right]$$

and similarly $f_{i+1} - g_i \in \left(0, \frac{1}{(n+1)b_{i+1}} \right]$. We deduce that

$$f_1 < g_1 < f_2 < g_2 < f_3 < \dots < f_{d-1} < g_{d-1} < f_d.$$

Since x lies in one of the intervals $[f_i, g_i]$ or $[g_i, f_{i+1}]$ (for some i), the result follows. \square

Remark 3.69. If $\frac{1}{b(n+1)}$ is replaced with $\frac{1}{bn}$, a simple proof of the previous corollary, based uniquely on the pigeonhole principle, goes as follows: consider all numbers of the form $1 + [kx] - kx$ for $0 \leq k \leq n$. We have $n + 1$ numbers that belong to $[0, 1)$, so by the pigeonhole principle two must lie in an interval $(\frac{j}{n}, \frac{j+1}{n}]$ for some $0 \leq j < n$. We deduce the existence of integers u_1, u_2 and $0 \leq v_1 < v_2 \leq n$ such that

$$|u_2 - u_1 - x(v_2 - v_1)| < \frac{1}{n}.$$

Setting

$$b = \frac{v_2 - v_1}{\gcd(v_2 - v_1, u_2 - u_1)}, \quad a = \frac{u_2 - u_1}{\gcd(v_2 - v_1, u_2 - u_1)}$$

yields the desired result.

We are now ready to deal with the equation $x^2 + y^2 = n$. More precisely, we will prove the following beautiful theorem.

Theorem 3.70. *The map sending a pair (x, y) to $yx^{-1} \pmod{n}$ (where x^{-1} is the inverse of x modulo n) establishes a bijection between the set of pairs (x, y) of relatively prime positive integers x, y such that $x^2 + y^2 = n$ and the set of solutions of the congruence $z^2 \equiv -1 \pmod{n}$.*

Proof. Clearly if x, y are relatively prime positive integers such that $x^2 + y^2 = n$, then $\gcd(x, n) = 1$ (any common divisor of x and n would divide y^2 , but $\gcd(x, y^2) = 1$, so this divisor must be 1 or -1) and letting $z = yx^{-1} \pmod{n}$ we have

$$0 \equiv x^2 + y^2 \equiv x^2(z^2 + 1) \pmod{n},$$

thus $z^2 \equiv -1 \pmod{n}$ by Gauss' lemma. This shows that the map is well-defined.

Let us prove first the injectivity of the map. Consider two different pairs (x_1, y_1) and (x_2, y_2) that have the same image, say z . Thus $y_2 \equiv x_2 z \pmod{n}$

and $y_1 \equiv x_1 z \pmod{n}$. It follows that $x_1 y_2 \equiv x_2 y_1 \pmod{n}$, thus n divides $x_1 y_2 - x_2 y_1$. On the other hand

$$n^2 = (x_1^2 + y_1^2)(x_2^2 + y_2^2) = (x_1 y_2 - x_2 y_1)^2 + (x_1 y_1 + x_2 y_2)^2,$$

thus $|x_1 y_2 - x_2 y_1| < n$. We conclude that $x_1 y_2 = x_2 y_1$. But then Gauss' lemma yields $x_1 \mid x_2$ and $x_2 \mid x_1$, thus $x_1 = x_2$ and then $y_1 = y_2$, a contradiction.

Finally, let us deal with the most difficult part, the surjectivity of the map. Consider a positive integer z such that $z^2 \equiv -1 \pmod{n}$. We want to show that we can find relatively prime positive integers x, y such that $y \equiv xz \pmod{n}$ and $x^2 + y^2 = n$. By corollary 3.68 we can find relatively prime integers a, b such that $0 < b \leq \lfloor \sqrt{n} \rfloor$ and

$$\left| \frac{-z}{n} - \frac{a}{b} \right| \leq \frac{1}{b(1 + \lfloor \sqrt{n} \rfloor)} < \frac{1}{b\sqrt{n}}.$$

It follows that

$$0 < b^2 + (bz + an)^2 < n + n = 2n$$

and on the other hand

$$b^2 + (bz + an)^2 \equiv b^2 + b^2 z^2 = b^2(1 + z^2) \equiv 0 \pmod{n}.$$

Thus we must have

$$n = b^2 + (bz + an)^2.$$

In particular

$$b^2 \cdot \frac{z^2 + 1}{n} + 2abz - 1 + a^2 n = 0$$

and so $\gcd(b, n) = 1$ and also $\gcd(b, bz + an) = \gcd(b, an) = 1$. We deduce that if $bz + an > 0$ then $x = b$ and $y = bz + an$ work, while if $bz + an < 0$ then $x = -bz - an$ and $y = b$ work. The result follows. \square

We turn now to the diophantine equation $x^2 - dy^2 = 1$, where $d > 1$ is not a square (if d is a square, say $d = e^2$, then the equation can be written as $(x - ey)(x + ey) = 1$, the resolution being therefore very easy). This equation is widely known as Pell's equation, even though Pell did not have

major contributions to its study. Note that while studying Pell's equation we may assume that x and y are nonnegative. There is a trivial solution $(1, 0)$, but it is not clear at all that there are other solutions. We will prove now that there are infinitely many solutions of this equation. This requires a few preliminary steps.

We fix a positive integer d which is not a perfect square, so that \sqrt{d} is an irrational number by example 3.62.

Proposition 3.71. *There are infinitely many pairs (x, y) of positive integers such that*

$$|x - y\sqrt{d}| < \frac{1}{y}.$$

Proof. By corollary 3.68 for any $n \geq 1$ we can find integers $0 < b_n \leq n$ and a_n such that

$$|b_n\sqrt{d} - a_n| \leq \frac{1}{n+1} < \frac{1}{b_n}.$$

Note that necessarily $a_n > 0$. If the sequence $(b_n)_n$ has infinitely many distinct terms, we are done, so assume that this is not the case. Then the sequence $(a_n)_n$ is bounded and so has only finitely many distinct terms. It follows that there are indices i, j such that $b_n = b_i$ and $a_n = a_j$ for infinitely many n . But for such n we have

$$|b_i\sqrt{d} - a_i| \leq \frac{1}{n+1}$$

and the quantity $\frac{1}{n+1}$ becomes smaller (for n large enough) than any given positive real number. We deduce that $b_i\sqrt{d} = a_i$, contradicting the fact that \sqrt{d} is irrational. \square

The proof of the previous proposition adapts immediately to prove the following more general (and very useful) result:

Theorem 3.72. *If x is an irrational number, then there are infinitely many pairs (p, q) of integers with $\gcd(p, q) = 1$ and*

$$\left| x - \frac{p}{q} \right| < \frac{1}{q^2}.$$

On the other hand, the result established in the previous theorem fails for rational numbers, as the following exercise shows.

Example 3.73. Prove that if x is a rational number, then there are only finitely many rational numbers $\frac{p}{q}$ such that $|x - \frac{p}{q}| < \frac{1}{q^2}$.

Proof. Suppose that $\frac{p_n}{q_n}$ is an infinite sequence of pairwise distinct rational numbers (in lowest form) such that

$$|x - \frac{p_n}{q_n}| < \frac{1}{q_n^2}$$

for all n . If the sequence $(q_n)_{n \geq 0}$ is bounded, then the previous inequality shows that the sequence $(p_n)_{n \geq 0}$ is also bounded, which is clearly impossible. Writing $x = \frac{u}{v}$ in lowest form, we deduce that for infinitely many n we have $|q_n| > |v|$. But since by assumption

$$|q_n u - p_n v| < \frac{|v|}{|q_n|} < 1,$$

it follows that for infinitely many n we have $q_n u = p_n v$, contradicting the fact that the numbers $\frac{p_n}{q_n}$ are pairwise distinct. The result follows. \square

Theorem 3.74. Let d be a positive integer which is not a perfect square. The equation $x^2 - dy^2 = 1$ has integral solutions x, y with $x, y > 0$.

Proof. We will prove this in two steps. We first establish the existence of a nonzero integer k such that the equation $x^2 - dy^2 = k$ has infinitely many integral solutions with $x, y > 0$. Note that if x, y are positive integers and

$$|x - y\sqrt{d}| < \frac{1}{y},$$

then $x < y\sqrt{d} + 1 < 2y\sqrt{d}$ and so

$$|x^2 - dy^2| < \frac{1}{y}(x + y\sqrt{d}) < \frac{1}{y} \cdot 3y\sqrt{d} = 3\sqrt{d}.$$

Using proposition 3.71, we conclude that for infinitely many pairs (x, y) of positive integers we have $x^2 - dy^2 \in \{-N, \dots, -1, 1, \dots, N\}$ for some fixed positive integer $N = 1 + \lfloor 3\sqrt{d} \rfloor$. The result follows then from the pigeonhole principle.

Fix now a nonzero integer k such that $x^2 - dy^2 = k$ has infinitely many integral solutions with $x, y > 0$. Considering the pairs $(x \pmod{k}, y \pmod{k})$ for these solutions, we see (using again the pigeonhole principle) that we can find two solutions (x_1, y_1) and (x_2, y_2) for which $x_1 \equiv x_2 \pmod{k}$ and $y_1 \equiv y_2 \pmod{k}$. Setting

$$x = \frac{x_1x_2 - dy_1y_2}{k}, \quad y = \frac{x_1y_2 - x_2y_1}{k},$$

a simple calculation shows that

$$x^2 - dy^2 = \frac{1}{k^2}(x_1^2 - dy_1^2)(x_2^2 - dy_2^2) = 1.$$

On the other hand, since $x_1 \equiv x_2 \pmod{k}$ and $y_1 \equiv y_2 \pmod{k}$, we have

$$x_1x_2 - dy_1y_2 \equiv x_1^2 - dy_1^2 \equiv 0 \pmod{k}$$

and so x is an integer. Similarly y is an integer. If we prove that $y \neq 0$, then we are done (as then considering the numbers $|x|, |y|$ finishes the proof).

Assume now that $y = 0$, so that $x_1y_2 = x_2y_1$ and $x^2 = 1$ (since $x^2 - dy^2 = 1$). Thus $x = \pm 1$, i.e. $x_1x_2 - dy_1y_2 = \pm k$. Replacing x_2 by $\frac{x_1y_2}{y_1}$ yields

$$y_2(x_1^2 - dy_1^2) = \pm k \cdot y_1$$

and so $y_2 = \pm y_1$. Finally, we obtain $y_1 = y_2$ and $x_1 = x_2$, a contradiction. \square

We are now ready to express all positive solutions of the equation $x^2 - dy^2 = 1$ in terms of a distinguished solution. Namely, considering all pairs of positive integers (x, y) which satisfy $x^2 - dy^2 = 1$, it is clear that there is a unique pair (x, y) for which y has the smallest possible value (or equivalently $x + y\sqrt{d}$ has the smallest possible value). We call this pair the smallest positive solution. This solution generates all positive solutions, as the following theorem shows.

Theorem 3.75. *Let (x_1, y_1) be the smallest positive solution of the equation $x^2 - dy^2 = 1$. The general solution (x_n, y_n) is given by*

$$x_n + y_n\sqrt{d} = (x_1 + y_1\sqrt{d})^n.$$

We have

$$x_{n+1} = x_1x_n + dy_1y_n, \quad y_{n+1} = y_1x_n + x_1y_n$$

and the explicit formulae

$$x_n = \frac{(x_1 + y_1\sqrt{d})^n + (x_1 - y_1\sqrt{d})^n}{2}, \quad y_n = \frac{(x_1 + y_1\sqrt{d})^n - (x_1 - y_1\sqrt{d})^n}{2\sqrt{d}}.$$

Proof. Note that by the binomial formula and the fact that \sqrt{d} is irrational there are unique integers x_n, y_n such that

$$x_n + y_n\sqrt{d} = (x_1 + y_1\sqrt{d})^n,$$

and moreover they satisfy

$$x_n - y_n\sqrt{d} = (x_1 - y_1\sqrt{d})^n,$$

i.e. they are given by the explicit formulae appearing in the theorem. One sees directly that

$$x_{n+1} = x_1x_n + dy_1y_n, \quad y_{n+1} = y_1x_n + x_1y_n.$$

We have

$$x_n^2 - dy_n^2 = (x_n + y_n\sqrt{d}) \cdot (x_n - y_n\sqrt{d}) = (x_1^2 - dy_1^2)^n = 1,$$

thus (x_n, y_n) is a positive solution of the equation $x^2 - dy^2 = 1$. Conversely, consider a positive solution (x, y) of this equation and set

$$z_1 = x_1 + y_1\sqrt{d}, \quad z = x + y\sqrt{d}.$$

By minimality of z_1 , we have $z \geq z_1$. Since $z_1 > 1$, we deduce that there is a unique $n \geq 1$ such that

$$z_1^n \leq z < z_1^{n+1}.$$

Write

$$\frac{z}{z_1^n} = (x + y\sqrt{d})(x_1 - y_1\sqrt{d})^n = u + v\sqrt{d}$$

for some integers u, v , then $1 \leq u + v\sqrt{d} < z_1$. Note that, by the binomial formula and the fact that \sqrt{d} is irrational we also have

$$u - v\sqrt{d} = (x - y\sqrt{d})(x_1 + y_1\sqrt{d})^n$$

and so

$$u^2 - dv^2 = (x^2 - dy^2)(x_1^2 - dy_1^2)^n = 1.$$

Assuming that $u + v\sqrt{d} > 1$, we conclude that (u, v) is a positive² solution of the equation $x^2 - dy^2 = 1$ which is smaller than (x_1, y_1) , a contradiction. Thus $u + v\sqrt{d} = 1$ and so $z = z_1^n$, as needed. \square

Example 3.76. Are there integers $a, b > 1$ such that $ab + 1$ and $ab^3 + 1$ are both perfect squares?

Proof. Assume that such integers exist and write

$$ab + 1 = c^2, \quad ab^3 + 1 = x^2.$$

Then

$$x^2 - 1 = (c^2 - 1)b^2.$$

Consider this as a Pell equation in the variables x and b . Then smallest positive solution is obviously $x = c$ and $b = 1$, thus the general solution is given by the previous theorem. In particular, defining sequences x_n and b_n by

$$x_{n+1} = cx_n + (c^2 - 1)b_n, \quad b_{n+1} = x_n + cb_n$$

we deduce that $b = b_n$ for some n . Since $b > 1$, we must have $n > 1$. If $n \geq 3$, then $b_n \geq b_3 > c^2 - 1$, contradicting the fact that $b = b_n \mid c^2 - 1 = ab$. Thus $n = 2$ and $b = 2c$. It follows that $2c \mid c^2 - 1$, which is clearly impossible for $c > 1$. Thus there are no such a, b . \square

²To see that $u, v \geq 0$ note that $u - v\sqrt{d} = \frac{1}{u+v\sqrt{d}} \in (0, 1)$ since $u + v\sqrt{d} > 1$.

Finally, let us deal with the more general equation

$$ax^2 - by^2 = 1.$$

Example 3.77. Let a, b be positive integers such that $ab > 1$ is a square. Prove that the equation $ax^2 - by^2 = 1$ has no solutions in positive integers.

Proof. The existence of a solution (x, y) clearly forces $\gcd(a, b) = 1$. Since ab is a square, it follows that a and b are squares. Thus ax^2 and by^2 are consecutive positive perfect squares, which is impossible. \square

Example 3.78. Prove that there are no positive integers a, b such that $2a^2 + 1, 2b^2 + 1, 2(ab)^2 + 1$ are all perfect squares.

Proof. We argue by contradiction and assume that there are such integers. Clearly $a, b > 1$ and by symmetry we may assume that $a \geq b$. Then

$$4(2a^2 + 1)(2a^2b^2 + 1) = (4a^2b + b)^2 + 8a^2 - b^2 + 4$$

is a perfect square. However we clearly have

$$(4a^2b + b)^2 < (4a^2b + b)^2 + 8a^2 - b^2 + 4 < (4a^2b + b + 1)^2 = (4a^2b + b)^2 + 8a^2b + 2b + 1.$$

The result follows. \square

Theorem 3.79. Let a, b be positive integers such that $ab > 1$ is not a square. Let (x_1, y_1) be the smallest positive solution of the equation $ax^2 - by^2 = 1$ and let (u_n, v_n) be the general positive solution of the equation $u^2 - abv^2 = 1$. Then the general solution of the equation $ax^2 - by^2 = 1$ is given by (x_n, y_n) , with

$$x_n = x_1u_n + by_1v_n, \quad y_n = y_1u_n + ax_1v_n.$$

Proof. One checks that if (x, y) is a solution of the equation $ax^2 - by^2 = 1$, then $u = ax_1x - by_1y$ and $v = y_1x - x_1y$ is a solution of the equation $u^2 - abv^2 = 1$, and we can recover x and y from u and v by the formulae

$$x = x_1u + by_1v, \quad y = y_1u + ax_1v. \quad \square$$

Example 3.80. Let d be a positive integer which is not a perfect square and such that $x^2 - dy^2 = -1$ is solvable in positive integers. Let (x_0, y_0) be the smallest positive solution and define (x_1, y_1) by

$$x_1 + y_1\sqrt{d} = (x_0 + y_0\sqrt{d})^2.$$

Prove that (x_1, y_1) is the smallest positive solution of the equation $x^2 - dy^2 = 1$.

Proof. Clearly (x_1, y_1) is a positive solution of the equation $x^2 - dy^2 = 1$. Let (x_2, y_2) be the smallest positive solution of the equation $x^2 - dy^2 = 1$. For $i = 0, 1, 2$ set

$$z_i = x_i + y_i\sqrt{d}.$$

We prove first that $z_0 < z_2$. Assuming that $z_0 \geq z_2$, we clearly have $z_0 > z_2$ and so letting u, v be the integers such that

$$u + v\sqrt{d} = \frac{z_0}{z_2} = (x_0 + y_0\sqrt{d})(x_2 - y_2\sqrt{d})$$

we have

$$u^2 - dv^2 = (x_0^2 - dy_0^2)(x_2^2 - dy_2^2) = -1$$

and $u + v\sqrt{d} > 1$, as well as $u + v\sqrt{d} < z_0$, contradicting the minimality of (x_0, y_0) . Thus $z_0 < z_2$.

Assume next that $z_0^2 > z_2$, so letting u, v be integers with

$$u + v\sqrt{d} = \frac{z_2}{z_0} = (x_2 + y_2\sqrt{d})(x_0 - y_0\sqrt{d})$$

we obtain $u^2 - dv^2 = -1$ and $x_0 + y_0\sqrt{d} > u + v\sqrt{d} > 1$, contradicting again the minimality of (x_0, y_0) . Thus $z_0^2 \leq z_2$. Finally, by minimality of (x_2, y_2) it is clear that $z_0^2 \geq z_2$, thus $z_2 = z_0^2$ and we are done. \square

We deduce from the previous example that if the equation $x^2 - dy^2 = -1$ (with $d > 0$ not perfect square) has solutions in positive integers and (x_0, y_0) is the smallest positive solution, then all solutions in positive integers are given by considering the odd (and positive) powers of $x_0 + y_0\sqrt{d}$. Also, all solutions in positive integers of the equation $x^2 - dy^2 = 1$ are obtained by considering the even (and positive) powers of $x_0 + y_0\sqrt{d}$.

Example 3.81. Find all m, n positive integers such that $3^m = 2n^2 + 1$.

Proof. The answer is $(m, n) = (1, 1), (2, 2)$, and $(5, 11)$.

There are two cases to consider:

1) If m is even then $(3^{m/2}, n)$ forms a solution to $x^2 - 2y^2 = 1$. The solutions for x in this Pell equation are given by the recurrence formula

$$x_0 = 1, x_1 = 3, x_k = 6x_{k-1} - x_{k-2}.$$

It is easy to check that $3^2 = 9$ divides x_k if and only if $k \equiv 3 \pmod{6}$. But for such k , x_k is also divisible by 11, implying that $m/2$ does not exceed 1. Hence, $m = 2$ gives the only positive solution $(m, n) = (2, 2)$.

2) If m is odd then $(3^{(m-1)/2}, n)$ forms a solution to $3x^2 - 2y^2 = 1$. The solutions for x in this Pell-like equation are given by the recurrence formula

$$x_0 = 1, x_1 = 9, x_k = 10x_{k-1} - x_{k-2}.$$

It is easy to check that $3^3 = 27$ divides x_k iff $k \equiv 4 \pmod{9}$. But for such k , x_k is also divisible by 17, implying that $(m-1)/2$ does not exceed 2. Hence, $m = 1$ and $m = 5$ give the only solutions, $(m, n) = (1, 1)$ and $(5, 11)$. \square

Example 3.82. (Romania TST 2011) Prove that there are infinitely many positive integer numbers n such that $n^2 + 1$ has two positive divisors whose difference is n .

Proof. In formulas, we are asked to show that there are infinitely many solutions to $n^2 + 1 = d(n + d)$ in positive integers. This equation is equivalent to $(2d - n)^2 - 5n^2 = 4$. The Pell equation $x^2 - 5y^2 = 1$ has infinitely many solutions, and setting $n = 2y$ and $d = x + y$ gives infinitely many solutions to the desired equation. \square

Example 3.83. (AMM 10622) Find infinitely many triples (a, b, c) of positive integers forming an arithmetic progression and such that $ab + 1, bc + 1, ca + 1$ are all perfect squares.

Proof. Consider solutions (x, y) in positive integers of the Pell equation $x^2 - 3y^2 = 1$ and set

$$a = 2y - x, \quad b = 2y, \quad c = 2y + x.$$

Then

$$ab + 1 = 4y^2 - 2xy + 1 = y^2 - 2xy + x^2 = (y - x)^2, \quad bc + 1 = 4y^2 + 2xy + 1 = (y + x)^2$$

and

$$ca + 1 = 4y^2 - x^2 + 1 = y^2.$$

Since a, b, c clearly form an arithmetic progression and since the Pell equation above has infinitely many solutions in positive integers, the problem is solved. \square

Remark 3.84. One can prove (not without effort) that there are no positive integers a, b, c, d in arithmetic progression such that $ab + 1, ac + 1, ad + 1, bc + 1, bd + 1, cd + 1$ are all perfect squares.

Example 3.85. (AMM 10220) Let $x > 0$ be a real number. A positive integer n is x -suarish if one can write $n = ab$ for some integers a, b such that $1 \leq a \leq b < (1 + x)a$. Prove that there are infinitely many sequences of 6 consecutive positive integers in which each term is x -suarish.

Proof. We will try to impose that each of the numbers $n^2, n^2 - 1, n^2 - 2, n^2 - 3, n^2 - 4, n^2 - 5$ is x -suarish. Clearly $n^2, n^2 - 1 = (n - 1)(n + 1), n^2 - 4 = (n - 2)(n + 2)$ are x -suarish for n large enough, so it remains to deal with $n^2 - 2, n^2 - 3$ and $n^2 - 5$. Choosing n of the form $n = a^2 + a - 2$ for some integer $a > 1$, one checks that

$$n^2 - 2 = (n - a)(n + a + 1)$$

is x -suarish (if a is big enough) and so is

$$n^2 - 5 = (n - 2a + 1)(n + 2a + 3).$$

Finally, if we can also ensure that such n 's are of the form $n = 2b^2 - 2$ for some integer b , then

$$n^2 - 3 = (n - 2b + 1)(n + 2b + 1)$$

is also x -squarish for b big enough. It is thus sufficient to prove that for infinitely many positive integers a we can find integers b such that

$$a^2 + a - 2 = 2b^2 - 2.$$

This reduces to $(2a + 1)^2 - 8b^2 = 1$. Since the equation $u^2 - 8v^2 = 1$ has infinitely many positive solutions and u is odd in any such solution, the result follows. \square

Example 3.86. (AMM 10238) a) Prove that $1 + a$ and $1 + 3a$ are both perfect squares for infinitely many positive integers a .

b) Let $a_1 < a_2 < \dots$ be all positive integers satisfying the conditions of part a). Prove that $1 + a_n a_{n+1}$ is a perfect square for all n .

Proof. Imposing $1 + a = x^2$ and $1 + 3a = y^2$, we are reduced to showing that the Pell-like equation $y^2 - 3x^2 = -2$ has infinitely many positive solutions. Taking into account part b), we also need to find explicitly all solutions. For this, we observe first that for any solution (x, y) both x and y are odd (by taking the equation $y^2 - 3x^2 = -2$ modulo 4). Letting

$$u = \frac{3x - y}{2}, \quad v = \frac{y - x}{2}$$

we obtain positive integers u, v such that $u^2 - 3v^2 = 1$. The smallest positive solution of this last equation being $(2, 1)$, we deduce that all solutions are given by (u_n, v_n) , where

$$u_n + v_n\sqrt{3} = (2 + \sqrt{3})^n,$$

in other words

$$u_n = \frac{A^n + B^n}{2}, \quad v_n = \frac{A^n - B^n}{2\sqrt{3}},$$

where $A = 2 + \sqrt{3}$ and $B = 2 - \sqrt{3}$. Since we can recover x, y from u, v via $x = u + v$ and $y = 3v + u$, we deduce that

$$a_n = (u_n + v_n)^2 - 1,$$

with the notation introduced in part b) of the problem. This immediately implies part a). A simple but tedious computation yields then

$$1 + a_n a_{n+1} = \left(\frac{A^{2n+2} + B^{2n+2} - 8}{6} \right)^2.$$

It suffices therefore to prove that $\frac{A^{2n+2} + B^{2n+2} - 8}{6} = \frac{u_{2n+1} - 2}{3}$ is an integer for all n . This follows easily since the formula for u_n coming from the binomial formula applied to $(2 + \sqrt{3})^n$ shows $u_n \equiv 2^n \pmod{3}$. \square

Example 3.87. Solve in integers the equation

$$(x^2 - 1)(y^2 - 1) = \left(\left(\frac{x - y}{2} \right)^2 - 1 \right)^2.$$

Proof. Write the equation as

$$(xy)^2 - x^2 - y^2 + 1 = 1 - \frac{(x - y)^2}{2} + \left(\frac{x - y}{2} \right)^4$$

or equivalently as

$$\left(\left(\frac{x - y}{2} \right)^2 + xy \right) \cdot \left(\left(\frac{x - y}{2} \right)^2 - xy \right) + \frac{(x + y)^2}{2} = 0$$

and finally

$$\frac{(x + y)^2}{4} \cdot \frac{x^2 - 6xy + y^2}{4} + \frac{(x + y)^2}{2} = 0.$$

Thus either $x + y = 0$, giving the family of solutions $(t, -t)$, with $t \in \mathbf{Z}$, or $x^2 - 6xy + y^2 + 8 = 0$. This last equation is equivalent to $(y - 3x)^2 = 8(x^2 - 1)$. Hence $y - 3x$ is a multiple of 4, say $y - 3x = 4z$, and $x^2 - 2z^2 = 1$. Let (u_n, v_n) be the family of solutions of the Pell equation $u^2 - 2v^2 = 1$ in positive integers u and v . Then we get a second family of solutions with $x = \pm u_n$ and $z = \pm v_n$. Noting that $3u_n \pm 4v_n = u_{n \pm 1}$, we see that $y = \pm u_{n \pm 1}$. Being careful with the signs, we find solutions $(x_n, y_n) = (u_n, u_{n+1})$, (u_{n+1}, u_n) , $(-u_n, -u_{n+1})$, and $(-u_{n+1}, -u_n)$. \square

Example 3.88. Prove that the only positive integers n for which $3^n - 2$ is a perfect square are $n = 1$ and $n = 3$.

Proof. Suppose that $n > 3$ is a solution of the problem. Write $u^2 = 3^n - 2$ and observe that n must be odd, since otherwise the right-hand side is congruent to $-1 \pmod{8}$, and no square is congruent to $-1 \pmod{8}$. Let $v = 3^{\frac{n-1}{2}}$, so that $u^2 - 3v^2 = -2$. As we saw in example 3.2.4, this Pell-like equation can be reduced to a Pell equation by analyzing parities. Letting (u_n, v_n) the general positive solution of this equation, with $u_0 = v_0 = 1$, we have $v_1 = 3$ and $v_{n+2} = 4v_{n+1} - v_n$ for all $n \geq 0$. It is not difficult (though rather tedious) to check that v_n is a multiple of 9 if and only if $n \equiv 4 \pmod{9}$ and that v_n is a multiple of 17 if and only if $n \equiv 4 \pmod{9}$. Writing $v = 3^{\frac{n-1}{2}} = v_k$ for some k , we have $9 \mid v = v_k$ since $n > 3$. We deduce that $n \equiv 4 \pmod{9}$ and so $17 \mid v_k$. Since v_k is a power of 3, this is clearly impossible. Hence any solution n satisfies $n \leq 3$, and the result follows easily. \square

Example 3.89. (USA TST 2013) Determine if there exists a (three-variable) polynomial $P(x, y, z)$ with integer coefficients and with the following property: a positive integer n is not a perfect square if and only if there is a triple (x, y, z) of positive integers such that $P(x, y, z) = n$.

Proof. We will prove that there is such a polynomial $P \in \mathbf{Z}[X, Y, Z]$, more precisely that the polynomial

$$P(X, Y, Z) = Z^2(X^2 - ZY^2 - 1)^2 + Z$$

is a solution of the problem.

If n is not a perfect square, then the Pell equation $x^2 - ny^2 = 1$ has nontrivial solutions. Choosing $z = n$ yields $P(x, y, n) = n$. On the other hand, suppose that $P(x, y, z) = n$ for some triple (x, y, z) of positive integers. Then

$$z^2(x^2 - zy^2 - 1)^2 + z = n.$$

Assume that n is a perfect square, then $x^2 - zy^2 - 1$ is nonzero and

$$(z(x^2 - zy^2 - 1))^2 < n < (z|x^2 - zy^2 - 1| + 1)^2,$$

a contradiction. \square

Example 3.90. (Putnam 2000) Prove that for infinitely many positive integers n each of the numbers $n, n+1$ and $n+2$ can be written as the sum of two squares of integers.

Proof. Choosing n of the form $n = x^2 - 1$ for some $x > 1$, the numbers $n+1 = x^2 + 0^2$ and $n+2 = x^2 + 1^2$ are automatically sums of two squares. It remains to ensure that n itself is a sum of two squares for suitable x . Simply choose x such that $x^2 - 2y^2 = 1$ for some $y > 1$. As this Pell-type equation has infinitely many solutions, we are done. \square

Remark 3.91. One can avoid the use of Pell's equation here, by choosing $x = 2y^2 + 1$ for some $y > 0$, in which case

$$n = x^2 - 1 = (2y)^2 + (2y^2)^2.$$

3.3 Least common multiple

In this section we study the dual notion of gcd, namely that of least common multiple. We will see very soon that the two notions are closely linked to each other.

Definition 3.92. Let a_1, \dots, a_n be nonzero integers, not all equal to 0. The least common multiple of a_1, \dots, a_n , denoted $\text{lcm}(a_1, \dots, a_n)$ is the smallest positive integer which is divisible by a_1, a_2, \dots, a_n .

Note that the definition makes sense: the set of positive integers divisible by a_1, \dots, a_n is nonempty, since $|a_1 \dots a_n|$ is such an integer (as a_1, \dots, a_n are nonzero). We make the convention that $\text{lcm}(a_1, \dots, a_n) = 0$ when one of the a_i 's is equal to 0.

Before moving on to theoretical properties of the lcm function, let us mention the following beautiful problem of Erdős:

Example 3.93. Let n be an integer greater than 1. Integers $1 < a_1 < a_2 < \dots < a_k < n$ have the property that $\text{lcm}(a_i, a_j) > n$ for all $1 \leq i \neq j \leq k$. Prove that:

$$\frac{1}{a_1} + \frac{1}{a_2} + \dots + \frac{1}{a_k} < \frac{3}{2}.$$

Proof. The idea is extremely beautiful: let us count the number of multiples of one of the numbers a_1, \dots, a_k in the set $\{1, 2, \dots, n\}$. For each $1 \leq i \leq k$ there are $\left\lfloor \frac{n}{a_i} \right\rfloor$ such multiples of a_i . The crucial claim is that no multiple of a_i can be equal to a multiple of a_j for some $1 \leq i \neq j \leq k$, since this common multiple would be at least $\text{lcm}(a_i, a_j) > n$. Thus the total number of multiples of one of the numbers a_1, \dots, a_k is $\sum_{i=1}^k \left\lfloor \frac{n}{a_i} \right\rfloor$, and in particular (using that $\lfloor x \rfloor > x - 1$)

$$n \geq \sum_{i=1}^k \left\lfloor \frac{n}{a_i} \right\rfloor > \sum_{i=1}^k \frac{n}{a_i} - k.$$

The problem is then reduced to proving that $k \leq \frac{n}{2}$. But if $k > \frac{n}{2}$ then by Erdős' problem 3.93 there are indices $i < j$ such that $a_i \mid a_j$, and so $a_j = \text{lcm}(a_i, a_j) > n$, a contradiction. \square

The following theorem is dual to the statement that any common divisor of a_1, \dots, a_n is a divisor of $\text{gcd}(a_1, \dots, a_n)$:

Theorem 3.94. *Any common multiple of the integers a_1, \dots, a_n is a multiple of $\text{lcm}(a_1, \dots, a_n)$.*

Proof. Let $l = \text{lcm}(a_1, \dots, a_n)$. We may assume that $l \neq 0$ (i.e. that all a_i are nonzero). Let x be a common multiple of a_1, \dots, a_n and assume that l does not divide x . Thus we can write $x = ql + r$ for integers q, r such that $0 < r < l$. But then $r = x - ql$ is a common multiple of a_1, \dots, a_n (since so are x and ql), and $0 < r < l$, contradicting the minimality of l . The result follows. \square

Example 3.95. Show that

$$\text{lcm}(1, 2, \dots, 2n) = \text{lcm}(n+1, n+2, \dots, 2n).$$

Proof. Let A denote the left-hand side and B the right-hand side. Since A is a multiple of $n+1, n+2, \dots, 2n$, and B is the smallest multiple of these numbers, we have $B \leq A$. To prove that $A \leq B$, it suffices to prove that B is a multiple of $1, 2, \dots, 2n$ and this reduces to checking that B is a multiple of $1, 2, \dots, n$. Fix $k \in \{1, 2, \dots, n\}$. Among the k consecutive numbers $n+1, n+2, \dots, n+k \leq 2n$ there is a multiple of k , and this multiple is a divisor of B by definition. Thus $k \mid B$ and the result follows. \square

Using the previous theorem, the reader can easily check that

$$\text{lcm}(a_1, \dots, a_n) = \text{lcm}(\text{lcm}(a_1, \dots, a_{n-1}), a_n)$$

for all integers a_1, \dots, a_n . Thus computing the lcm of a family of integers reduces to understanding the computation for two integers.

The link between the gcd and the lcm of two numbers is given by the following important result.

Theorem 3.96. *If a, b are positive integers, then*

$$\text{lcm}(a, b) = \frac{ab}{\text{gcd}(a, b)},$$

in other words

$$\text{lcm}(a, b) \cdot \text{gcd}(a, b) = ab.$$

In particular, if $\text{gcd}(a, b) = 1$ then $\text{lcm}(a, b) = ab$.

Proof. Let $d = \text{gcd}(a, b)$ and write $a = da_1$ and $b = db_1$ with $\text{gcd}(a_1, b_1) = 1$. By definition, $\text{lcm}(a, b) = dk$ for some integer k , and dk is a multiple of both a and b , thus k is a multiple of a_1 and b_1 . Since $\text{gcd}(a_1, b_1) = 1$, we deduce that $a_1b_1 \mid k$ and so $da_1b_1 = \frac{ab}{d}$ divides $\text{lcm}(a, b)$. Since on the other hand $\frac{ab}{d} = da_1b_1$ is a common multiple of a and b , we have $\frac{ab}{d} \geq \text{lcm}(a, b)$. Thus $\text{lcm}(a, b) = \frac{ab}{d}$ and the theorem is proved. \square

Let us mention the following useful consequence of theorem 3.96:

Corollary 3.97. *For all integers $0 < a < b$ we have*

$$\text{lcm}(a, b) \geq \frac{ab}{b-a},$$

or equivalently

$$\frac{1}{\text{lcm}(a, b)} \leq \frac{1}{a} - \frac{1}{b}.$$

Proof. It suffices to observe that $\text{gcd}(a, b)$ is a positive divisor of $b - a$, thus $\text{gcd}(a, b) \leq b - a$. The result follows. \square

Here is a beautiful and rather classical application of the previous corollary:

Example 3.98. (Kvant M 865) Prove that for any integers $1 \leq a_0 < a_1 < \dots < a_n$ we have

$$\sum_{k=0}^{n-1} \frac{1}{\text{lcm}(a_k, a_{k+1})} \leq 1 - \frac{1}{2^n}.$$

Proof. We will prove this inequality by induction, the case $n = 1$ being clear. Suppose now that the inequality is true for any choice of $1 \leq a_0 < \dots < a_{n-1}$ and fix $1 \leq a_0 < \dots < a_n$.

Using corollary 3.97, we obtain

$$\sum_{k=0}^{n-1} \frac{1}{\text{lcm}(a_k, a_{k+1})} \leq \sum_{k=0}^{n-1} \left(\frac{1}{a_k} - \frac{1}{a_{k+1}} \right) = \frac{1}{a_0} - \frac{1}{a_n} \leq 1 - \frac{1}{a_n},$$

the inequality holds if $a_n \leq 2^n$.

Suppose now that $a_n > 2^n$, hence $\text{lcm}(a_{n-1}, a_n) \geq a_n > 2^n$. Using the inductive hypothesis, we obtain

$$\sum_{k=0}^{n-1} \frac{1}{\text{lcm}(a_k, a_{k+1})} < 1 - \frac{1}{2^{n-1}} + \frac{1}{2^n} = 1 - \frac{1}{2^n},$$

proving the inductive step in this case also. □

We continue with a few other illustrations of theorem 3.96.

Example 3.99. (Kvant) Let a and b be positive integers such that

$$\frac{\text{lcm}(a, b)}{\text{gcd}(a, b)} = a - b.$$

Prove that $\text{lcm}(a, b) = (\text{gcd}(a, b))^2$.

Proof. Set $d = \text{gcd}(a, b)$. Then $a = a_1 d, b = b_1 d$, where $\text{gcd}(a_1, b_1) = 1$. On the other hand

$$\text{lcm}(a, b) = \frac{ab}{\text{gcd}(a, b)} = \frac{d^2 a_1 b_1}{d} = d a_1 b_1.$$

Hence the given identity can be written as $a_1 b_1 = d(a_1 - b_1)$. It follows that a_1 divides db_1 , i.e. a_1 divides d . Analogously b_1 divides d . Thus $a_1 b_1$ divides d and we conclude that $a_1 - b_1 = 1$ and $d = a_1 b_1$. Hence

$$\text{lcm}(a, b) = da_1 b_1 = d^2 = (\gcd(a, b))^2.$$

Remark. The above arguments show that all positive integers a and b satisfying the given identity have the form

$$a = n(1 + n)^2, \quad b = (1 + n)n^2,$$

where n is a positive integer. □

Example 3.100. (Saint Peterburg 2009) Let x, y, z be pairwise different positive integers such that

$$\text{lcm}(x, y) - \text{lcm}(x, z) = y - z.$$

Prove that x divides both y and z .

Proof. Since the left-hand side is a multiple of x , so must be the right-hand side, thus we can write $y - z = kx$ for some integer k . Then

$$\text{lcm}(x, y) = \frac{xy}{\gcd(x, y)} = \frac{xy}{\gcd(x, z + kx)} = \frac{xy}{\gcd(x, z)}.$$

Since $\text{lcm}(x, z) = \frac{xz}{\gcd(x, z)}$, we can rewrite the equation as

$$\frac{x(y - z)}{\gcd(x, z)} = y - z.$$

Since $y \neq z$, we deduce that $x = \gcd(x, z)$ and so $x \mid z$. Since $x \mid y - z$, we conclude that $x \mid y$, too. □

Example 3.101. (Romania JBMO TST 2007) Find all positive integers n which can be written as $\text{lcm}(a, b) + \text{lcm}(b, c) + \text{lcm}(c, a)$ for some positive integers a, b, c .

Proof. Call such integers good. Clearly, if n is good, then so is $2n$ (simply replace a, b, c by $2a, 2b$ and $2c$). By choosing $b = c = 1$ we see that all odd integers greater than 1 are good, hence by the first observation all integers except for powers of 2 are good. Now, we will prove that powers of 2 are bad, thus finishing the solution.

Suppose that $2^k = \text{lcm}(a, b) + \text{lcm}(b, c) + \text{lcm}(c, a)$ for some positive integers a, b, c . Clearly $k > 1$. We may write $a = 2^A a_1, b = 2^B b_1, c = 2^C c_1$ with $A \geq B \geq C$ (by symmetry we may assume this) and a_1, b_1, c_1 odd. We deduce that

$$2^k = 2^A(\text{lcm}(a_1, b_1) + \text{lcm}(a_1, c_1)) + 2^B \text{lcm}(b_1, c_1).$$

Dividing by 2^B we obtain a power of 2 greater than 1 in the left-hand side and an odd number in the right-hand side (note that $\text{lcm}(a_1, b_1) + \text{lcm}(a_1, c_1)$ is even), which is clearly absurd. \square

Algebraic identities can be very powerful when trying to understand the lcm of a family of numbers a_1, \dots, a_n . The idea is the following: one tries to find integers b_1, \dots, b_n such that one has total control on $\frac{b_1}{a_1} + \dots + \frac{b_n}{a_n}$. Since this expression is clearly of the form $\frac{k}{\text{lcm}(a_1, \dots, a_n)}$ for some integer k , this leads to nontrivial information about $\text{lcm}(a_1, \dots, a_n)$ (such as order of growth or divisibility properties). Combinatorial identities are very powerful in finding suitable b_1, \dots, b_n as above. So are techniques coming from algebra, such as the Lagrange's interpolation formula, which leads to numerous algebraic identities. Let us recall this last result. Consider pairwise distinct real numbers a_1, \dots, a_n and arbitrary real numbers b_1, \dots, b_n . Then Lagrange's interpolation polynomial

$$P(X) = \sum_{k=1}^n b_k \prod_{j \neq k} \frac{X - a_j}{a_k - a_j}$$

is the unique polynomial of degree $\leq n-1$ such that $P(a_k) = b_k$ for $1 \leq k \leq n$. Indeed, it is easy to see that this polynomial satisfies $P(a_k) = b_k$ for $1 \leq k \leq n$ (since $\prod_{j \neq k} \frac{X - a_j}{a_k - a_j}$ vanishes at a_j for any $j \neq k$). If Q is another polynomial of degree $\leq n-1$ such that $Q(a_k) = b_k$ for $1 \leq k \leq n$, then $P - Q$ has degree $\leq n-1$ and at least n different roots (namely a_1, \dots, a_n), thus must be the zero polynomial, which gives $P = Q$.

Let us give a few examples of how algebraic identities can be used to obtain interesting properties of $\text{lcm}(a_1, \dots, a_n)$.

Example 3.102. Let $a > b \geq n$ be positive integers. Prove that

$$\text{lcm}(1, 2, \dots, n) \cdot \frac{\binom{a}{n} - \binom{b}{n}}{a - b} \in \mathbf{Z}.$$

Proof. Let $k = a - b$ and let $N = \text{lcm}(1, 2, \dots, n)$. We need to prove that $\frac{N}{k}(\binom{b+k}{n} - \binom{b}{n})$ is an integer. Using Vandermonde's identity

$$\binom{a+b}{n} = \sum_{k=0}^n \binom{a}{k} \binom{b}{n-k}$$

we obtain

$$\frac{N}{k} \left(\binom{b+k}{n} - \binom{b}{n} \right) = \frac{N}{k} \sum_{i=1}^n \binom{b}{n-i} \binom{k}{i},$$

so it is enough to check that $\frac{N}{k} \binom{k}{i}$ is an integer for $1 \leq i \leq n$. But

$$\frac{N}{k} \binom{k}{i} = \frac{N}{i} \cdot \binom{k-1}{i-1}$$

and $\frac{N}{i}$ is an integer by the definition of N . □

Example 3.103. Prove that for all positive integers a, b we have

$$a \binom{a+b}{b} \mid \text{lcm}(b+1, b+2, \dots, b+a).$$

Proof. Using a partial fraction decomposition of $\frac{n!}{(x+1)(x+2)\dots(x+n)}$, we obtain the identity

$$\frac{n!}{(x+1)(x+2)\dots(x+n)} = \sum_{i=1}^n \frac{(-1)^{i-1} i \binom{n}{i}}{x+i}.$$

Therefore

$$\frac{1}{\binom{a+b}{b}} = \frac{a!}{(b+1)\dots(b+a)} = \sum_{i=1}^a \frac{(-1)^{i-1} i \binom{a}{i}}{b+i} = a \sum_{i=1}^a \frac{(-1)^{i-1} \binom{a-1}{i-1}}{b+i}.$$

The last expression is clearly of the form $\frac{ak}{\text{lcm}(b+1, \dots, b+a)}$ for some integer k . Thus $a \binom{a+b}{b} \mid \text{lcm}(b+1, b+2, \dots, b+a)$, as needed. \square

Example 3.104. Prove that for all integers $n > 1$ we have

$$(n+1) \text{lcm} \left(\binom{n}{0}, \binom{n}{1}, \dots, \binom{n}{n} \right) = \text{lcm}(1, 2, \dots, n+1).$$

Proof. Let

$$N = \text{lcm}(1, 2, \dots, n+1).$$

First, we prove that the left-hand side divides N . It suffices to prove that

$$(n+1) \binom{n}{i} = (i+1) \binom{n+1}{i+1}$$

divides N for all $0 \leq i \leq n$, which follows directly from the previous example. On the other hand, we claim that the left-hand side is a multiple of N . For this, it suffices to prove that $i+1$ divides the lcm of the numbers $(n+1) \binom{n}{j} = (j+1) \binom{n+1}{j+1}$, which is clear since $i+1$ divides $(i+1) \binom{n+1}{i+1}$. The result follows. \square

Example 3.105. Prove that for all $n > 1$

$$\text{lcm}(1, 2, \dots, n) \geq 2^{n-1}.$$

Proof. This follows directly from the previous example, since

$$\text{lcm} \left(\binom{n}{0}, \binom{n}{1}, \dots, \binom{n}{n} \right) \geq \frac{1}{n+1} \sum_{i=0}^n \binom{n}{i} = \frac{2^n}{n+1}. \quad \square$$

Example 3.106. (Saint Petersburg 2004) Given distinct positives a_1, a_2, \dots, a_n . Let $b_i = (a_i - a_1)(a_i - a_2) \dots (a_i - a_{i-1})(a_i - a_{i+1}) \dots (a_i - a_n)$. Prove that $\text{lcm}(b_1, b_2, \dots, b_n)$ is divisible by $(n-1)!$.

Proof. For any polynomial f of degree $< n$ we have

$$f(X) = \sum_{k=1}^n f(a_k) \prod_{j \neq k} \frac{X - a_j}{a_k - a_j}.$$

In particular, for any $f(X) = cX^{n-1} + dX^{n-2} + \dots$ we have (by looking at the coefficients of X^{n-1})

$$c = \sum_{k=1}^n \frac{f(a_k)}{\prod_{j \neq k} (a_k - a_j)} = \sum_{k=1}^n \frac{f(a_k)}{b_k}.$$

If moreover $f(a_k)$ is an integer for all $1 \leq k \leq n$, then the expression $\sum_{k=1}^n \frac{f(a_k)}{b_k}$ is clearly of the form $\frac{u}{\text{lcm}(b_1, \dots, b_n)}$ for some integer u , in particular $\text{lcm}(b_1, \dots, b_n) \cdot c$ is an integer. Take now

$$f(X) = \binom{X}{n-1} = \frac{1}{(n-1)!} X(X-1)\dots(X-n+2).$$

In this case $c = \frac{1}{(n-1)!}$ and $f(a_k) = \binom{a_k}{n-1}$ is an integer for all $1 \leq k \leq n$. We deduce that $(n-1)!$ divides $\text{lcm}(b_1, \dots, b_n)$. \square

3.4 Problems for practice

Bézout's theorem and Gauss' lemma

1. Prove that for all positive integers a, b, c we have

$$\gcd(a, bc) \mid \gcd(a, b) \cdot \gcd(a, c).$$

2. (Romania TST 1990) Let a, b be relatively prime positive integers. Let x, y be nonnegative integers and let n be a positive integer for which

$$ax + by = a^n + b^n.$$

Prove that

$$\left\lfloor \frac{x}{b} \right\rfloor + \left\lfloor \frac{y}{a} \right\rfloor = \left\lfloor \frac{a^{n-1}}{b} \right\rfloor + \left\lfloor \frac{b^{n-1}}{a} \right\rfloor.$$

3. (Kvant M 1996) Find all $n > 1$ for which there exist pairwise different positive integers a_1, a_2, \dots, a_n such that

$$\frac{a_1}{a_2} + \frac{a_2}{a_3} + \dots + \frac{a_{n-1}}{a_n} + \frac{a_n}{a_1}$$

is an integer.

4. Let m, n be positive integers greater than 1. We define the sets

$$P_m = \left\{ \frac{1}{m}, \frac{2}{m}, \dots, \frac{m-1}{m} \right\} \text{ and } P_n = \left\{ \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n} \right\}.$$

Find

$$\min\{|a - b| : a \in P_m, b \in P_n\}.$$

5. (Saint Petersburg 2004) Positive integers m, n, k are such that $5^n - 2$ and $2^k - 5$ are multiples of $5^m - 2^m$. Prove that $\gcd(m, n) = 1$.
6. (Russia 2000) Sasha tries to find a positive integer $X \leq 100$. He can choose any two positive integers M, N less than 100 and ask for $\gcd(X + M, N)$. Prove that he can find X after 7 questions.
7. (Poland 2002) Let k be a fixed positive integer. The sequence $\{a_n\}_{n \geq 1}$ is defined by

$$a_1 = k + 1, a_{n+1} = a_n^2 - ka_n + k.$$

Show that if $m \neq n$, then the numbers a_m and a_n are relatively prime.

8. (Romania TST 2005) Let m, n be relatively prime positive integers with m even and n odd. Prove that

$$\sum_{k=1}^{n-1} (-1)^{\lfloor \frac{mk}{n} \rfloor} \left\{ \frac{mk}{n} \right\} = \frac{1}{2} - \frac{1}{2n}.$$

We denoted by $\{x\}$ the fractional part of x , i.e. $\{x\} = x - \lfloor x \rfloor$.

9. An infinite sequence a_1, a_2, \dots of positive integers has the property that $\gcd(a_m, a_n) = \gcd(m, n)$ for all $m \neq n \geq 1$. Prove that $a_n = n$ for all $n \geq 1$.

10. (Iran 2011) Prove that there are infinitely many positive integers n such that $n^2 + 1$ has no proper divisor of the form $k^2 + 1$.
11. a) (Romanian Masters in Mathematics 2009) Let a_1, \dots, a_k be nonnegative integers and let $d = \gcd(a_1, \dots, a_k)$ and $n = a_1 + \dots + a_k$. Prove that

$$\frac{d}{n} \cdot \frac{n!}{a_1! \dots a_k!} \in \mathbf{Z}.$$

- b) Prove that $(n)!^k k! | (nk)!$ for all positive integers n, k .
12. (Brazil 2011) Are there 2011 positive integers $a_1 < a_2 < \dots < a_{2011}$ such that $\gcd(a_i, a_j) = a_j - a_i$ for any i, j such that $1 \leq i < j \leq 2011$?
13. (Tournament of the Towns 2001) Are there positive integers $a_1 < a_2 < \dots < a_{100}$ such that

$$\gcd(a_1, a_2) > \gcd(a_2, a_3) > \dots > \gcd(a_{99}, a_{100}) > \gcd(a_{100}, a_1)?$$

14. (Russian Olympiad 2012) Let n be an integer greater than 1. When a runs over all integers greater than 1, what is the maximum number of pairwise relatively prime numbers among $1 + a, 1 + a^2, \dots, 1 + a^{2^n - 1}$?
15. (Brazilian Olympic Revenge 2014) a) Prove that for all positive integers n we have

$$\gcd\left(n, \left\lfloor n\sqrt{2} \right\rfloor\right) < \sqrt[4]{8n^2}.$$

- b) Prove that there are infinitely many positive integers n such that

$$\gcd\left(n, \left\lfloor n\sqrt{2} \right\rfloor\right) > \sqrt[4]{7.99n^2}.$$

16. (AMM) The greatest common divisor of a set D of positive integers is 1. Prove the existence of a bijection $f : \mathbf{Z} \rightarrow \mathbf{Z}$ such that $|f(n) - f(n-1)| \in D$ for all integers n .

17. (China TST 2012) Let n be an integer greater than 1. Prove that there are only finitely many n -tuples of positive integers (a_1, a_2, \dots, a_n) such that
- a) $a_1 > a_2 > \dots > a_n$ and $\gcd(a_1, a_2, \dots, a_n) = 1$.
- b) We have

$$a_1 = \gcd(a_1, a_2) + \gcd(a_2, a_3) + \dots + \gcd(a_{n-1}, a_n) + \gcd(a_n, a_1).$$

Applications to diophantine equations and approximations

18. Integers a, b and rational numbers x, y satisfy $y^2 = x^3 + ax + b$. Prove that we can write $x = \frac{u}{v^2}$ and $y = \frac{w}{v^3}$ for some integers u, v, w , with $\gcd(u, v) = \gcd(w, v) = 1$.
19. (Kvant M 905) Let x and n be positive integers such that $4x^n + (x+1)^2$ is a perfect square. Prove that $n = 2$ and find at least one x with this property.
20. Solve in positive integers the equation

$$\frac{1}{x^2} + \frac{1}{y^2} = \frac{1}{z^2}.$$

21. (Romania TST 2015) A Pythagorean triple is a solution (x, y, z) of the equation $x^2 + y^2 = z^2$ in positive integers, where we count (x, y, z) and (y, x, z) as the same triple. Given a non-negative integer n , prove that some positive integer appears in precisely n distinct Pythagorean triples.
22. Find all triples (x, y, n) of positive integers with $\gcd(x, n+1) = 1$ and $x^n + 1 = y^{n+1}$.
23. Let n be a positive integer such that n^2 is the difference of the cubes of two consecutive positive integers. Prove that n is the sum of the squares of two consecutive positive integers.

24. (Vietnam 2007) Let x, y be integers different from -1 such that $\frac{x^4-1}{y+1} + \frac{y^4-1}{x+1}$ is also an integer. Prove that $x^4y^{44} - 1$ is a multiple of $x + 1$.
25. (Balkan 2006) Find all triplets of positive rational numbers (m, n, p) such that the numbers $m + \frac{1}{np}$, $n + \frac{1}{pm}$, $p + \frac{1}{mn}$ are all integers.
26. A polynomial f has integer coefficients and satisfies $|f(a)| = |f(b)| = 1$ for some distinct integers a, b .
- a) Prove that if $|a - b| > 2$, then f has no rational root.
- b) Prove that if $|a - b| = 2$, then the only possible rational root of f is $\frac{a+b}{2}$.
27. (Turkey 2003) Find all positive integers n for which $2^{2n+1} + 2^n + 1$ is a perfect power.
28. Let f be a polynomial with rational coefficients such that for all positive integers n the equation $f(x) = n$ has at least one rational solution. Prove that $\deg(f) = 1$.

Least common multiple

29. (Kyiv mathematical festival 2014)
- a) Let y be a positive integer. Prove that for infinitely many positive integers x we have

$$\text{lcm}(x, y+1) \cdot \text{lcm}(x+1, y) = x(x+1).$$

- b) Prove that there exists positive integer y such that

$$\text{lcm}(x, y+1) \cdot \text{lcm}(x+1, y) = y(y+1)$$

for at least 2014 positive integers x .

30. (Kvant M 666) Find the least positive integer a for which there exist pairwise different positive integers a_1, a_2, \dots, a_9 greater than a such that

$$\text{lcm}(a, a_1, a_2, \dots, a_9) = 10a.$$

31. (Korea 2013) Find all functions $f : \mathbf{N} \rightarrow \mathbf{N}$ satisfying

$$f(mn) = \text{lcm}(m, n) \cdot \text{gcd}(f(m), f(n))$$

for all positive integers m, n .

32. (Romania TST 1995) Let $f(n) = \text{lcm}(1, 2, \dots, n)$. Prove that for any $n \geq 2$ one can find a positive integer x such that

$$f(x) = f(x+1) = \dots = f(x+n).$$

33. Prove that for all positive integers a_1, \dots, a_n

$$\text{lcm}(a_1, \dots, a_n) \geq \frac{a_1 a_2 \dots a_n}{\prod_{1 \leq i < j \leq n} \text{gcd}(a_i, a_j)}.$$

34. (AMM 3834) Let $n > 4$ and let $a_1 < a_2 < \dots < a_n \leq 2n$ be positive integers. Prove that

$$\min_{1 \leq i \neq j \leq n} \text{lcm}(a_i, a_j) \leq 6(\lfloor n/2 \rfloor + 1).$$

35. Let $(a_n)_{n \geq 1}$ be a sequence of integers such that $m - n \mid a_m - a_n$ for all $m, n \geq 1$. Suppose that there is a polynomial f such that $|a_n| \leq f(n)$ for all $n \geq 1$. Prove that there is a polynomial P with rational coefficients such that $a_n = P(n)$ for all $n \geq 1$.

36. Let n, k be positive integers and let $1 < a_1 < \dots < a_k \leq n$ be a sequence of integers such that $\text{lcm}(a_i, a_j) \leq n$ for all $1 \leq i, j \leq k$. Prove that $k \leq 2 \lfloor \sqrt{n} \rfloor$.

37. (AMM E 3350) For $n \geq 1$ and $1 \leq k \leq n$ define

$$A(n, k) = \text{lcm}(n, n-1, \dots, n-k+1).$$

Let $f(n)$ be the largest k such that $A(n, 1) < A(n, 2) < \dots < A(n, k)$.

- Prove that $f(n) \leq 3\sqrt{n}$.
- Prove that $f(n) > k$ if $n > k! + k$.

38. Let $a_1 < a_2 < \dots < a_n$ be an arithmetic progression of positive integers such that a_1 is relatively prime to the common difference. Prove that $a_1 a_2 \dots a_n$ divides $(n-1)! \cdot \text{lcm}(a_1, \dots, a_n)$.
39. Let $n > 1$ and let $a_0 < a_1 < \dots < a_n$ be positive integers such that $\frac{1}{a_0}, \dots, \frac{1}{a_n}$ is an arithmetic progression. Prove that

$$a_0 \geq \frac{2^n}{n+1}.$$

Chapter 4

The fundamental theorem of arithmetic

This chapter is devoted to the proof and the many consequences of the fundamental theorem of arithmetic: the unique factorization of integers into products of prime numbers. Basic properties of prime and composite numbers are studied, with many examples. These are then applied to prove the fundamental theorem of arithmetic, and the remaining part of this chapter is devoted to applications of this theorem, for instance to the study of arithmetic functions.

4.1 Composite numbers

We start by defining prime and composite numbers. Prime numbers are the bricks of arithmetic, and most of the material in this book will be devoted to a better understanding of this notion.

Definition 4.1. a) An integer $n > 1$ is called a prime number (or simply prime) if the only positive divisors of n are 1 and n , in other words if n has no proper divisors.

b) An integer $n > 1$ is called composite if it is not a prime number, in other words if there is an integer $1 < d < n$ such that $d \mid n$, or equivalently if

$n = ab$ for some integers $a, b > 1$.

Note that even though the only positive divisor of 1 is 1, but we do not consider 1 to be a prime. There are many reasons for this. For example, if 1 were called prime, then the unique factorization of integers into products of prime numbers would need a cumbersome restatement. The sequence of primes starts as

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, \dots$$

It is not clear for now that there are infinitely many prime numbers, but we will prove later on that this is indeed the case.

Before focusing on primes, let us spend some time dealing with composite numbers. First of all, note that there are many composite numbers: all even integers greater than 2 are composite, and also all multiples of 3 greater than 3, all multiples of 4, etc. It looks therefore natural to conjecture that most of the integers greater than 1 are composite: for instance, if n is large enough, then more than 99.99999 percent of the integers between 1 and n are composite. Though this looks intuitively true, the proof of this statement is already nontrivial and we will be able to prove it only after having introduced a fair amount of theory.

Since we are dealing with the basics for now, we can only prove the following weak result, which is already very important historically: prime numbers have unbounded gaps, that is for any N there are two consecutive primes whose difference is greater than N . Establishing that there are infinitely many pairs of consecutive prime numbers that have bounded difference is a much deeper problem and was only established in the spectacular work of Yitang Zhang in 2013: he showed that there are infinitely many pairs of consecutive primes which differ by at most $70 \cdot 10^6$. This was later improved in several articles to 270. Replacing 270 with 2 and therefore proving the famous twin primes conjecture (saying that there are infinitely many primes p such that $p + 2$ is a prime) will probably require a great deal of new ideas. The fact that primes have unbounded gaps is equivalent to the following:

Proposition 4.2. *For any $n > 1$ there are n consecutive composite numbers.*

Proof. The numbers $(n+1)!+2, (n+1)!+3, \dots, (n+1)!+n+1$ are n consecutive

composite numbers, since i divides $(n+1)!+i$ for $2 \leq i \leq n+1$ and $(n+1)!+i > i$. \square

Example 4.3. Is there a sequence of 2005 consecutive positive integers that contains exactly 25 primes?

Proof. The answer is positive. Let $f(n)$ be the number of primes among $n+1, n+2, \dots, n+2005$. One easily checks that $f(1) > 25$. The key observation is that $f(n+1) - f(n)$ is either $-1, 0$ or 1 . Indeed, if both $n+1$ and $n+2006$ are both composite or both prime, then $f(n+1) - f(n) = 0$. If only $n+1$ is prime, then $f(n+1) - f(n) = -1$ and if only $n+2006$ is prime, then $f(n+1) - f(n) = 1$. Since there are arbitrarily long strings of consecutive composite integers, there is n such that $f(n) = 0$. Since f cannot increase or decrease by more than 1 at a time, it follows that there must be k such that $f(k) = 25$. \square

The next example is a more elaborate version of the proof of proposition 4.2.

Example 4.4. (Kvant M 2284) Prove that there exists a strictly increasing sequence a_1, a_2, \dots of positive integers such that for any arithmetic progression b_1, b_2, \dots of positive integers all but finitely many terms of the sequence $a_1 + b_1, a_2 + b_2, \dots$ are composite.

Proof. We will show that the sequence $a_n = (n^2)!$, $n \geq 1$ has the desired property. Let b_1, b_2, \dots be an arithmetic progression of positive integers with common difference d , so that $b_k = b_1 + (k-1)d$. For $k \geq \max(b_1, d)$ we have $b_k \leq k \cdot \max(b_1, d) \leq k^2$, thus for $n > \max(b_1, d)$ the number $a_n + b_n$ is divisible by $b_n > 1$ and so it is not a prime. \square

The next example is also historically very important: it shows that in any nonconstant polynomial sequence there are infinitely many composite numbers. In other words, nonconstant polynomial sequences cannot generate only primes.

Theorem 4.5. (Goldbach) *Let f be a nonconstant polynomial with integer coefficients and with positive leading coefficient. There are infinitely many composite numbers in the sequence $f(1), f(2), f(3), \dots$*

Proof. Since f has positive leading coefficient, there is an n such that $f(n) > 1$. Note that $f(n + kf(n)) \equiv f(n) \pmod{kf(n)}$, thus $f(n) \mid f(n + kf(n))$ for all k . But $f(n + kf(n))$ is a nonconstant polynomial in k with positive leading coefficient. Hence there is a K such that for all $k \geq K$, we have $f(n + kf(n)) > f(n)$. Hence $f(n + kf(n))$ is composite for $k \geq K$. \square

Remark 4.6. If we consider polynomials in several variables, the situation can change rather drastically: Jones, Sato and Wada constructed a polynomial f in 26 variables a, b, c, \dots such that when a, b, c, \dots range over the nonnegative integers, the positive numbers among $f(a, b, c, \dots)$ are precisely the prime numbers!

In the next examples, we discuss a few methods that are often used to prove that a given number is composite. Algebraic identities can be used from time to time to establish that numbers are composite.

Example 4.7. (Komal A 622) Prove that $\frac{7^{7^n+1}+1}{7^{7^n}+1}$ is composite for all $n \geq 1$.

Proof. The key ingredient is the algebraic identity

$$\frac{x^7 + 1}{x + 1} = (x + 1)^6 - 7x(x^2 + x + 1)^2$$

Checking that this holds is a purely mechanical matter, which we will leave to the reader. It follows that if $x = 7y^2$ for some $y > 1$ (which is the case when $x = 7^{7^n}$ with $n \geq 1$) then

$$\frac{x^7 + 1}{x + 1} = ((x + 1)^3 - 7y(x^2 + x + 1))((x + 1)^3 + 7y(x^2 + x + 1)).$$

If we prove that $(x + 1)^3 - 7y(x^2 + x + 1)$ is greater than 1, then we can conclude that $\frac{x^7+1}{x+1}$ is composite. But

$$x^2 + x + 1 = \frac{x^3 - 1}{x - 1} < \frac{x^3}{x - 1}$$

and $(x + 1)^3 > x^3 + 1$, thus it suffices to check that $7y < x - 1$, or equivalently $7y^2 - 7y - 1 > 0$. This is clear for $y > 1$, so we are done. \square

Congruences are also a very useful tool in proving that a given integer is composite. Here are a few examples:

Example 4.8. Prove that $521 \cdot 12^n + 1$ is composite for all $n \geq 1$.

Proof. If n is odd, then $521 \cdot 12^n + 1 \equiv 521 \cdot (-1)^n + 1 \equiv 0 \pmod{13}$ and we are done. If $n \equiv 0 \pmod{4}$, we work mod 29 (since $12^2 + 1 = 5 \cdot 29$, hence $12^4 \equiv 1 \pmod{29}$) and get

$$521 \cdot 12^n + 1 \equiv 521 + 1 = 522 = 18 \cdot 29 \equiv 0 \pmod{29}.$$

Finally, if $n \equiv 2 \pmod{4}$, then $521 \cdot 12^n + 1 \equiv 2^n + 1 \equiv 2^2 + 1 \equiv 0 \pmod{5}$ and we are done again. \square

Remark 4.9. One can also prove that $78557 \cdot 2^n + 1$ is composite for all $n \geq 1$, by proving that it is a multiple of one of the numbers 3, 5, 7, 13, 19, 37 or 73. We do not know whether for any $a > 1$ there is $k > 0$ such that $k \cdot a^n + 1$ is composite for all n .

Example 4.10. (Kvant) The sequence of positive integers a_1, a_2, \dots satisfies $a_{n+2} = a_n a_{n+1} + 1$ for all $n \geq 1$. Prove that if $n \geq 9$ then $a_n - 22$ is composite.

Proof. Let $n \geq 1$ and let $k = a_{n+1}$. Then $a_{n+2} \equiv 1 \pmod{k}$, $a_{n+3} = a_{n+1}a_{n+2} + 1 \equiv 1 \pmod{k}$, $a_{n+4} = a_{n+2}a_{n+3} + 1 \equiv 2 \pmod{k}$ and similarly $a_{n+5} \equiv 3 \pmod{k}$, $a_{n+6} \equiv 7 \pmod{k}$ and $a_{n+7} \equiv 22 \pmod{k}$. Hence $k \mid a_{n+7} - 22$. In other words, $a_{n+1} \mid a_{n+7} - 22$ for all $n \geq 1$. We want to prove that $a_{n+7} - 22$ is composite for $n \geq 2$. Note that $a_1 \geq 1$, $a_2 \geq 1$ and the recurrence relation immediately yields $a_{n+6} \geq 21$. Moreover, the recurrence relation also gives $a_{n+5} \geq a_{n+1} + 1$. Thus $a_{n+7} = a_{n+5}a_{n+6} + 1 > a_{n+1} + 22$ and so $a_{n+7} - 22$ is composite. \square

Remark 4.11. The same proof shows that if $b_1 = 1, b_2 = 1$ and $b_{n+2} = b_n b_{n+1} + 1$, then $a_n - b_k$ is composite for $n \geq k + 3$, since it is a multiple of a_{n-k} greater than a_{n-k} .

Example 4.12. (Putnam, 2010) Prove that for each positive integer n , the number $10^{10^{10^n}} + 10^{10^n} + 10^n - 1$ is composite.

Proof. Put $N = 10^{10^{10^n}} + 10^{10^n} + 10^n - 1$. Write $n = 2^m k$ with m a nonnegative integer and k a positive odd integer. For each nonnegative integer j ,

$$10^{2^m j} \equiv (-1)^j \pmod{10^{2^m} + 1}.$$

Since $10^n \geq n \geq 2^m \geq m + 1$, 10^n is divisible by 2^{m+1} , and similarly 10^{10^n} is divisible by 2^{10^n} and hence by 2^{m+1} . It follows that

$$N \equiv 1 + 1 + (-1) + (-1) \equiv 0 \pmod{10^{2^m} + 1}.$$

Since $N \geq 10^{10^n} > 10^n + 1 \geq 10^{2^m} + 1$, it follows that N is composite. \square

4.2 The fundamental theorem of arithmetic

In this section we will prove the fundamental theorem of arithmetic: the existence and uniqueness of prime factorization for integers greater than 1. This theorem will be constantly used from now on.

4.2.1 The theorem and its first consequences

We start with a weak form, the existence of the factorization.

Theorem 4.13. *Any integer $n > 1$ is a product of (not necessarily distinct) prime numbers.*

Proof. We argue by contradiction and assume that $n > 1$ is the smallest counterexample. In particular, n is not a prime number, hence it must have a proper divisor d . Since n is the smallest counterexample, d and $\frac{n}{d}$ are the product of some prime numbers. But then $n = \frac{n}{d} \cdot d$ is also the product of some primes, contradiction. The result follows. \square

The uniqueness of prime factorization is deeper and relies on the following fundamental theorem, which establishes a crucial and not formal property of prime numbers. Despite the rather easy-looking statement, the next theorem is not at all a formal consequence of the definition of a prime and the proof requires Gauss' lemma (which required Bézout's theorem, which itself required the Euclidean division...). Fortunately, we have already done all the hard work.

Theorem 4.14. *Let a, b be integers and let p be a prime divisor of ab . Then $p \mid a$ or $p \mid b$.*

Proof. Suppose that p does not divide a . Then $\gcd(a, p) = 1$, since $\gcd(a, p)$ is a positive divisor of p and cannot be p . Since $p \mid ab$ and $\gcd(a, p) = 1$, Gauss' lemma yields $p \mid b$, finishing the proof. \square

A useful corollary (which will be considerably refined in later chapters) of the previous theorem is the following:

Corollary 4.15. *Let p be a prime and let a be an integer not divisible by p . There is a positive integer k such that $p \mid a^k - 1$.*

Proof. Consider the remainders of the numbers $1, a, a^2, \dots$ when divided by p . Since there are only finitely many remainders, the pigeonhole principle yields the existence of $0 \leq i < j$ such that a^i and a^j give the same remainder when divided by p . Thus $p \mid a^i(a^{j-i} - 1)$. Since p does not divide a , the previous theorem yields $p \mid a^{j-i} - 1$ and so we can take $k = j - i$. \square

We are now ready to state and prove the fundamental theorem of arithmetic:

Theorem 4.16. *(Fundamental theorem of arithmetic) Any integer $n > 1$ can be uniquely written as a product of prime numbers, up to the order of the factors. In other words, if p_1, p_2, \dots, p_k and q_1, \dots, q_l are prime numbers such that $p_1 p_2 \dots p_k = q_1 \dots q_l$ then $k = l$ and there is a permutation σ of $1, 2, \dots, k$ such that $q_i = p_{\sigma(i)}$ for $1 \leq i \leq k$.*

Proof. The existence has already been established. In order to prove uniqueness, it suffices to prove the statement concerning $p_1, \dots, p_k, q_1, \dots, q_l$. We will prove this by induction on $k + l$, using the previous theorem. The base case $k + l = 2$ is clear. Since p_1 divides $q_1 \dots q_l$, the previous theorem shows that there exists i such that p_1 divides q_i . Since p_1 and q_i are primes, this forces $p_1 = q_i$. By permuting q_1, \dots, q_l , we may assume that $i = 1$. Dividing by p_1 we obtain $p_2 \dots p_k = q_2 \dots q_l$ and the number of factors in the products decreases. Hence we can apply the inductive hypothesis to conclude. \square

If an integer $n > 1$ is a product of primes $p_1 p_2 \dots p_k$, we say that p_1, \dots, p_k are the prime divisors or prime factors of n . In other words, a prime p is a prime factor or prime divisor of n if $p \mid n$. Note that if $a, b > 1$ are integers, then the set of prime factors of ab is the union of the set of prime factors of a and that of b , since a prime p divides ab if and only if p divides a or p divides b .

By collecting equal numbers among p_1, \dots, p_k in the equality $n = p_1 p_2 \dots p_k$, we deduce that n can be written as

$$n = q_1^{\alpha_1} q_2^{\alpha_2} \dots q_s^{\alpha_s}$$

with q_1, \dots, q_s pairwise different prime numbers and $\alpha_1, \dots, \alpha_s$ positive integers. This is called the prime factorization (or canonical factorization) of n . Note that by the fundamental theorem of arithmetic the numbers q_1, \dots, q_s and $\alpha_1, \dots, \alpha_s$ are unique.

The fundamental theorem of arithmetic describes the multiplicative structure of the set of integers, in terms of prime numbers. The additive structure of the set of integers is relatively simple, but the interaction between the two structures is the source of many very difficult (and most of the time unsolved) problems. For instance, one of the oldest and most intractable problems (so far) is the famous Goldbach conjecture, stating that any even integer greater than 2 can be written as the sum of two prime numbers. A weaker version of this conjecture (also known as the ternary Goldbach problem) states that any odd number greater than 5 can be written as the sum of three (not necessarily distinct) primes. After almost one century of hard work (starting with Hardy and Littlewood in 1923, Vinogradov in 1937 and ending with Helfgott in 2013), this weaker conjecture is now a theorem.

Another famous conjecture relating the additive and multiplicative structure of integers was stated in 1986 by Masser and Oesterlé. In order to state it, let us introduce a notation: if n is an integer greater than 1, let

$$r(n) = \prod_{p \mid n} p$$

be the product of all different prime factors of n .

Conjecture 4.17. (*the abc conjecture*) For any $\varepsilon > 0$ there is a constant $c(\varepsilon) > 0$ such that for all nonzero integers a, b, c satisfying $a + b + c = 0$ and $\gcd(a, b, c) = 1$ we have

$$\max(|a|, |b|, |c|) < c(\varepsilon) \cdot r(abc)^{1+\varepsilon}.$$

This conjecture lies extremely deep, since it is not difficult to prove that it implies many difficult results, which are either already theorems or still conjectural. For instance, the abc conjecture immediately implies that Fermat's last theorem holds for all sufficiently large n : if n is large enough, then the equation $x^n + y^n = z^n$ has no integer solutions with $xyz \neq 0$. Indeed, suppose x, y, z is such a solution (with x, y, z positive to simplify notations). Then we may assume that $\gcd(x, y, z) = 1$ and hence

$$z^n < c(1/2)r(xyz)^{\frac{3}{2}} \leq c(1/2)z^{\frac{9}{2}}.$$

Since $z \geq 2$ (otherwise $xy = 0$) we deduce that

$$2^{n-\frac{9}{2}} < c(1/2),$$

which bounds n from above.

Similarly, it is a simple exercise to deduce from the abc conjecture the following result (which is a deep theorem of Darmon and Granville, proved independently of the abc conjecture): if $p, q, r \geq 2$ and the equation $x^p + y^q = z^r$ has infinitely many solutions in positive integers with $\gcd(x, y, z) = 1$, then

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} \geq 1.$$

Indeed, for any $\varepsilon > 0$ and any solution we will have

$$z^r \leq c(\varepsilon)r(xyz)^{1+\varepsilon} \leq c(\varepsilon)z^{(1+\varepsilon)\left(\frac{r}{p}+\frac{r}{q}+1\right)}.$$

We deduce that

$$1 \leq \left(\frac{1}{p} + \frac{1}{q} + \frac{1}{r}\right) \cdot (1 + \varepsilon)$$

and since $\varepsilon > 0$ was arbitrary, the result follows.

It is not difficult to check that the only triples (p, q, r) with $p, q, r \geq 2$ and

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} > 1$$

are $(2, 2, n)$ (with any $n \geq 2$), $(2, 3, 3)$, $(2, 3, 4)$ and $(2, 3, 5)$ and their permutations, while the only triples with

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$$

are $(3, 3, 3)$, $(2, 4, 4)$, $(2, 3, 6)$ and their permutations. For instance, we have already seen that the equation $x^4 + y^2 = z^4$ has no nontrivial solutions. On the other hand, one can prove (with a lot of work!) that the only nontrivial solutions of the equation $x^3 + y^6 = z^2$ are $2^3 + (\pm 1)^6 = (\pm 3)^2$. In a different direction, each of the equations $x^3 + y^3 = z^2$, $x^4 + y^3 = z^2$, $x^4 + y^2 = z^3$, $x^5 + y^3 = z^2$ have infinitely many solutions, for instance for the equation $x^3 + y^3 = z^2$ a family of solutions is given by

$$x = a^4 + 6a^2b^2 - 3b^4, y = -a^4 + 6a^2b^2 + 3b^4, z = 6ab(a^4 + 3b^4)$$

with arbitrary positive integers a, b . These are not the only solutions, for instance another infinite family of solutions is given by

$$x = a^4 + 8ab^3, y = -4a^3b + 4b^4, z = a^6 - 20a^3b^3 - 8b^6.$$

Yet more examples of nontrivial solutions of such equations are given by

$$9262^3 + 15312283^2 = 113^7, 33^8 + 1549034^2 = 15613^3, 3^5 + 11^4 = 122^2, \dots$$

The remaining part of this section is devoted to a long series of exercises and examples that illustrate the previous fundamental theoretical results.

Example 4.18. (Zhautykov Olympiad 2010) Find all primes p, q such that

$$p^3 - q^7 = p - q.$$

Proof. Write the equation as

$$\begin{aligned} p(p^2 - 1) &= q(q^6 - 1) = q(q^3 - 1)(q^3 + 1) \\ &= q(q - 1)(q^2 + q + 1)(q + 1)(q^2 - q + 1). \end{aligned}$$

Thus p divides one of the numbers $q, q - 1, q^2 + q + 1, q + 1, q^2 - q + 1$. We claim that $p > q^2$, which then implies that $p = q^2 + q + 1$. Indeed, if $p \leq q^2$, then

$$q(q^6 - 1) = p^3 - p < p^3 - 1 \leq q^6 - 1,$$

impossible.

Hence $p = q^2 + q + 1$ and the equation becomes

$$p^2 - 1 = q(q - 1)(q + 1)(q^2 - q + 1)$$

or equivalently

$$q(q + 1)(q^2 + q + 2) = q(q + 1)(q - 1)(q^2 - q + 1).$$

Dividing by $q(q + 1)$ and simplifying the resulting expressions yields

$$(q - 3)(q^2 + 1) = 0,$$

hence $q = 3$ and then $p = 11$. □

Example 4.19. (Saint Petersburg 2013) Find all primes p, q such that $2p - 1$, $2q - 1$ and $2pq - 1$ are all perfect squares.

Proof. Say $2p - 1 = a^2$, $2q - 1 = b^2$ and $2pq - 1 = c^2$ for some positive integers a, b, c . Then $p \mid a^2 + 1$ and $p \mid c^2 + 1$, thus $p \mid c^2 - a^2$ and so $p \mid c - a$ or $p \mid c + a$. Note that a, c are odd and p is odd, hence $p \leq \frac{c+a}{2}$ and with a similar argument $q \leq \frac{b+c}{2}$. In other words

$$c \geq 2p - a, \quad c \geq 2q - b.$$

But then

$$2pq - 1 = c^2 \geq (2p - a)(2q - b) = 4pq - 2pb - 2qa + ab,$$

which becomes

$$2pq + 1 + ab \leq 2pb + 2qa.$$

In particular $pq < pb + qa$ and so

$$1 < \frac{b}{q} + \frac{a}{p} < \sqrt{\frac{2}{q}} + \sqrt{\frac{2}{p}}.$$

We may assume that $p \leq q$. The previous inequality yields $p \leq 7$. Clearly $p = 7$ is not a solution since $2p - 1$ is not a square in this case. Thus $p \leq 5$ and since $p = 2$ and $p = 3$ are not solutions, we obtain $p = 5$. But then

$$c = \sqrt{10q - 1} \geq 2q - b = 2q - \sqrt{2q - 1},$$

which immediately implies that $q \leq 5$ and then $q = 5$. Hence $p = q = 5$ is the only solution. \square

Next, we discuss a series of exercises in which theorem 4.14 is used to prove that certain numbers are composite.

Example 4.20. (Kvant M 888) Let a, b, c, d be positive integers such that $ab = cd$. Prove that for every positive integer k the number $a^k + b^k + c^k + d^k$ is composite.

Proof. Replacing a, b, c, d with a^k, b^k, c^k, d^k , we may assume that $k = 1$. Let us write $\frac{a}{c} = \frac{d}{b} = \frac{m}{n}$ in lowest terms, where m, n are positive integers. Since m divides na and $\gcd(m, n) = 1$, we must have $m \mid a$, hence $a = mu$ and $c = nu$ for some positive integer u . Similarly $d = mv$ and $b = nv$ for some positive integer v . But then

$$a + b + c + d = mu + nu + mv + nv = (m + n)(u + v)$$

is composite.

Here is another proof, more in the spirit of the argument used to solve the next exercise: assume that $a + b + c + d = p$ is a prime and note that $a + b \equiv -c - d \pmod{p}$ and $ab \equiv (-c) \cdot (-d) \pmod{p}$ (the first congruence is clear and the second one follows from the hypothesis of the problem). Thus

the coefficients of the polynomial $(X-a)(X-b) - (X+c)(X+d)$ are multiples of p and so its value at a is a multiple of p . In other words, $p \mid (a+c)(a+d)$. Since p is a prime, p divides one of the numbers $a+c$ and $a+d$, which is not the case since $p > \max(a+c, a+d)$. \square

Example 4.21. Let a, b, c, d be positive integers such that

$$a^2 + ab + b^2 = c^2 + cd + d^2.$$

Prove that $a + b + c + d$ is composite.

Proof. Assume that $p = a + b + c + d$ is a prime number. Then $a + b \equiv -c - d \pmod{p}$ hence

$$a^2 + b^2 + ab + ab \equiv c^2 + d^2 + cd + cd \pmod{p},$$

which combined with the hypothesis yields $ab \equiv cd \pmod{p}$. Considering the polynomial $(X-a)(X-b) - (X+c)(X+d)$ and arguing as in the previous example we deduce that $-(a+c)(a+d)$ is a multiple of p . It follows that p divides one of the numbers $a+c$ and $a+d$, which is impossible since p is greater than each of them. \square

Example 4.22. (IMO Shortlist 2005) Let a, b, c, d, e, f be positive integers such that $S = a+b+c+d+e+f$ divides $abc+def$ and $ab+bc+ca-de-ef-df$. Prove that S is composite.

Proof. Suppose that S is a prime and let $x = -d, y = -e, z = -f$, so that

$$a + b + c \equiv x + y + z \pmod{S}, \quad ab + bc + ca \equiv xy + yz + zx \pmod{S}$$

and $abc \equiv xyz \pmod{S}$. Considering the polynomial

$$(T-a)(T-b)(T-c) - (T-x)(T-y)(T-z)$$

and arguing as in the previous examples we obtain that

$$S \mid (a-x)(a-y)(a-z) = (a+d)(a+e)(a+f).$$

Since S is a prime, S divides one of the numbers $a+d, a+e, a+f$, which is impossible since S is greater than any of them. Hence S is composite. \square

Remark 4.23. There are many exercises (more or less difficult) with a very similar flavor and solution. Here are two more examples, left to the reader:

a) (IMO 2001) Let $a > b > c > d$ be positive integers such that

$$ac + bd = (b + d + a - c)(b + d - a + c).$$

Prove that $ab + cd$ is composite.

b) (USAMO 2015) Let a, b, c, d, e be distinct positive integers such that $a^4 + b^4 = c^4 + d^4 = e^5$. Prove that $ac + bd$ is a composite number.

Example 4.24. (IMO Shortlist 2001) Is it possible to find 100 positive integers not exceeding 25000, such that all pairwise sums of two of them are different?

Proof. We will prove more generally that for any odd prime p we can find $p-1$ numbers a_1, \dots, a_{p-1} not exceeding $2p^2$ and such that all pairwise sums of two of them are different (then taking $p = 101$ will solve the problem). If a is an integer, let \bar{a} be the remainder of a when divided by p . Let $a_n = 2np + \bar{n}^2$ for $1 \leq n \leq p-1$. The numbers a_1, \dots, a_{p-1} are smaller than $2p(p-1) + p < 2p^2$ and it remains to see that the pairwise sums are different. Suppose that $a_n + a_m = a_k + a_l$ for some $n \neq m$ and $k \neq l$ between 1 and $p-1$. We write this equality as

$$2p(n + m - k - l) = \bar{k}^2 + \bar{l}^2 - \bar{m}^2 - \bar{n}^2.$$

The right-hand side is between $2 - 2(p-1)$ and $2(p-1) - 2$ and is a multiple of $2p$ (since the left-hand side is so). Thus we must have $n + m = k + l$ and $\bar{k}^2 + \bar{l}^2 = \bar{m}^2 + \bar{n}^2$. We deduce that $n^2 + m^2 \equiv k^2 + l^2 \pmod{p}$. Combined with $n^2 + m^2 + 2mn = k^2 + l^2 + 2lk$ and using the fact that p is odd we obtain $nm \equiv lk \pmod{p}$. Thus the coefficients of the polynomial

$$(X - m)(X - n) - (X - l)(X - k)$$

are multiples of p and so $p \mid (m - l)(m - k)$. We deduce that either $m = l$ and $n = k$ or $m = k$ and $n = l$. \square

Recall that the Fibonacci sequence $(f_n)_{n \geq 1}$ is defined by $f_1 = f_2 = 1$ and $f_n = f_{n-1} + f_{n-2}$ for $n \geq 3$. It is not difficult to prove that if f_n is a prime

number, then n is also a prime number or $n = 4$, but the converse does not hold since f_{19} is not a prime. Many prime Fibonacci numbers are known (for instance one of the largest ones is $f_{1968721}$), but it is not known whether the Fibonacci sequence contains infinitely many primes. The following interesting result describes all primes in the shifted Fibonacci sequence $(f_n + 1)_{n \geq 1}$. A crucial ingredient in the proof is Catalan's identity

$$f_n^2 - f_{n+r}f_{n-r} = (-1)^{n-r}f_r^2,$$

whose proof is left as an easy exercise for the reader, recalling that we have the classical formula

$$f_n = \frac{\phi^n - (-\phi)^n}{\sqrt{5}}$$

where $\phi = \frac{1+\sqrt{5}}{2}$. Another crucial ingredient is theorem 4.14.

Example 4.25. a) Prove the Gelin-Cesàro identity

$$f_n^4 - f_{n-2}f_{n-1}f_{n+1}f_{n+2} = 1, \quad n \geq 3.$$

b) Find all n for which $f_n + 1$ is a prime.

Proof. a) We use Catalan's identity with $r = 1, 2$ to obtain

$$f_{n+1}f_{n-1} - f_n^2 = (-1)^n = f_n^2 - f_{n+2}f_{n-2}.$$

Thus $f_n^2 - 1$ and $f_n^2 + 1$ are $f_{n-1}f_{n+1}$ and $f_{n-2}f_{n+2}$ in some order. The desired result follows since in either case

$$f_n^4 - 1 = (f_n^2 - 1)(f_n^2 + 1) = f_{n-2}f_{n-1}f_{n+1}f_{n+2}.$$

b) It is easy to check that $f_n + 1$ is a prime for $n = 1, 2, 3$. Suppose that $n > 3$ and that $f_n + 1$ is a prime. Since $f_n + 1$ divides $f_n^4 - 1 = f_{n-2}f_{n-1}f_{n+1}f_{n+2}$, it must divide one of the numbers $f_{n-2}, f_{n-1}, f_{n+1}, f_{n+2}$. Since it is greater than f_{n-2} and f_{n-1} , it either divides f_{n+1} or f_{n+2} . On the other hand, it is clear that the Fibonacci sequence is increasing and so $f_{n+2} < 2f_{n+1}$ for all n , thus $f_{n+2} < 4f_n$. If $f_n + 1 \mid f_{n+1}$, since $f_{n+1} < 2(f_n + 1)$ we must have $f_{n+1} = f_n + 1$ and then $f_{n-1} = 1$, impossible for $n > 3$. Thus $f_n + 1 \mid f_{n+2}$ and

since $f_{n+2} < 4(f_n + 1)$ and $f_{n+2} > f_n + 1$ we must have $f_{n+2} = 2(f_n + 1)$ or $f_{n+2} = 3(f_n + 1)$. In the first case we obtain $f_{n+1} = f_n + 2$, that is $f_{n-1} = 2$, then $n = 4$ and $f_4 + 1 = 4$ is not a prime. In the second case we obtain $f_{n+1} = 2f_n + 3$, impossible since $f_{n+1} < 2f_n$. Thus the only solutions of the problem are $n = 1, 2, 3$. \square

4.2.2 The smallest and largest prime divisor

The next problems are concerned with the largest and smallest prime factors of integers. We will introduce therefore the following notation: if $n > 1$ is an integer, $P(n)$ will denote the largest prime factor of n , while $p(n)$ will denote the smallest prime factor of n .

The first two examples exploit a very specific property of monic quadratic polynomials. This class of polynomials can be characterized by the property that

$$f(X)f(X+1) = f(X+f(X)).$$

We leave it to the reader as a very pleasant exercise to prove this property. In particular, if $q(n) = P(f(n))$ is the largest prime factor of $f(n)$, then the previous relation yields

$$q(n+f(n)) = \max(q(n), q(n+1)).$$

Example 4.26. (IMO Shortlist 2013) Prove that there exist infinitely many positive integers n such that the largest prime divisor of $n^4 + n^2 + 1$ is equal to the largest prime divisor of $(n+1)^4 + (n+1)^2 + 1$.

Proof. Let $f(X) = X^2 - X + 1$, then $f(X+1) = X^2 + X + 1$ and so the previous identity becomes

$$f(n^2+1) = n^4 + n^2 + 1 = f(n)f(n+1), \quad (1)$$

Letting $q(n)$ be the largest prime factor of $f(n)$, the problem requires

$$q(n^2+1) = q((n+1)^2+1)$$

for infinitely many n , or equivalently (thanks to relation (1))

$$\max(q(n), q(n+1)) = \max(q(n+1), q(n+2))$$

for infinitely many n . This is the case if $q(n+1) \geq \max(q(n), q(n+2))$, and we will prove that this inequality holds for infinitely many n . Assume that this is not the case, hence $q(n+1) < \max(q(n), q(n+2))$ for $n \geq N$, where N is some positive integer. Since there is no infinite decreasing sequence of positive integers, there is $n_0 > N$ such that $q(n_0+1) \geq q(n_0)$. Since $q(n_0+1) < \max(q(n_0), q(n_0+2))$, we obtain $q(n_0+1) < q(n_0+2)$. Combining this with $q(n_0+2) < \max(q(n_0+1), q(n_0+3))$ we see that $q(n_0+2) < q(n_0+3)$ and inductively $q(n) < q(n+1)$ for $n > n_0$. But then the equality

$$q(n^2+1) = \max(q(n), q(n+1))$$

cannot hold for $n > n_0$, since $n^2 > n$ and $n^2 > n+1$. \square

Example 4.27. (Russia 2011) Let $q(n)$ be the largest prime divisor of n^2+1 . Prove that there are infinitely many pairwise distinct positive integers a, b, c such that $q(a) = q(b) = q(c)$.

Proof. Letting $f(X) = X^2+1$ we obtain

$$f(X^2+X+1) = f(X)f(X+1)$$

thus

$$q(n^2+n+1) = \max(q(n), q(n+1)) \quad (1)$$

and

$$q(n^2-n+1) = q((n-1)^2+(n-1)+1) = \max(q(n-1), q(n)).$$

Hence if $n > 1$ and $q(n) \geq \max(q(n-1), q(n+1))$ then

$$q(n) = q(n^2-n+1) = q(n^2+n+1)$$

and the numbers n, n^2-n+1 and n^2+n+1 are pairwise distinct. It suffices therefore to prove that for infinitely many n we have

$$q(n) \geq \max(q(n-1), q(n+1)),$$

which can be done exactly as in the previous problem. \square

The next problems deal with the smallest prime divisor of a number. Before discussing them, we would like to mention the following very important criterion of primality:

Proposition 4.28. *A number $n > 1$ is composite if and only if it has a prime divisor $p \leq \sqrt{n}$, that is if the smallest prime factor of n does not exceed \sqrt{n} .*

Proof. If n has such a prime factor, it is clear that n is composite. Conversely, suppose that n is composite, so we can write $n = ab$, with $a, b > 1$. Then each of a, b has at least one prime factor, say p and q . Since $n \geq pq$, we deduce that $\min(p, q)$ is a prime factor of n , not exceeding \sqrt{n} . \square

Example 4.29. (Kvant M 557) Prove that each set of n pairwise relatively prime numbers greater than 1 and less than $(2n - 1)^2$ contains at least one prime.

Proof. Suppose that the given numbers a_1, a_2, \dots, a_n are all composite. Denote by q_i the least prime divisor of $a_i, 1 \leq i \leq n$ and assume without loss of generality that $q_1 < \dots < q_n$ (note that the q_i 's are pairwise distinct as $\gcd(a_i, a_j) = 1$ for $i \neq j$). Thus $q_1 \geq 2, q_2 \geq 3$ and $q_{i+1} \geq q_i + 2$ for $i \geq 2$, which easily yields $q_n \geq 2n - 1$. But then $a_n \geq q_n^2 \geq (2n - 1)^2$, a contradiction. \square

Example 4.30. (Russia 2014) Find all integers $n > 1$ such that for any positive divisor a of n the number $a + 1$ divides $n + 1$.

Proof. Clearly all odd primes are solutions of the problem. Conversely, suppose that n is a solution and let us prove that n is a prime. If not, then n has a proper divisor $a \geq \sqrt{n}$ (namely n/p , where p is a prime factor $\leq \sqrt{n}$ of n). By hypothesis $a + 1$ divides $n + 1$, thus $a + 1$ divides $n + 1 - (a + 1) = n - a$. Since $a \mid n - a$ and since $\gcd(a, a + 1) = 1$, we deduce that $a(a + 1) \mid n - a$ and so $n - a \geq a(a + 1) > a^2 \geq n$, a contradiction. Thus the solutions of the problem are the odd prime numbers. \square

Example 4.31. (Saint Petersburg 2008) If a is an integer greater than 1, let $p(a)$ be its smallest prime factor. Let m, n be integers greater than 1.

a) Prove that if

$$m^2 + n = p(m) + p(n)^2$$

then $m = n$.

b) If

$$m + n = p(m)^2 - p(n)^2,$$

what are the possible values of m ?

Proof. a) We have $p(n)^2 - n = m^2 - p(m) > 0$, that is $p(n) > \sqrt{n}$. Thus n is a prime, say $n = q$, and $p(n) = q$. The equation becomes $m^2 - p(m) = q^2 - q$, hence $p(m)$ divides $q(q-1)$. Assume that $p(m)$ divides $q-1$, then $q > p(m)$ and $(m-q)(m+q) = p(m) - q < 0$, that is $m < q$. We conclude that

$$q^2 - q = m^2 - p(m) < m^2 \leq (q-1)^2,$$

a contradiction. Thus $p(m) = q$ and then $m^2 = q^2$, hence $m = q = n$.

b) We have this time

$$n + p(n)^2 = p(m)^2 - m$$

and again $p(m) > \sqrt{m}$, showing that m is a prime. If $m = 2$, then $n + p(n)^2 = 2$, which is impossible. Thus m is an odd prime. Conversely, if m is an odd prime, then we look for n even such that $n + p(n)^2 = p(m)^2 - m$, a relation which can also be written as $n + 4 = m^2 - m$. Thus $n = m^2 - m - 4$ works (note that $n > 1$ since $m \geq 3$). \square

Example 4.32. (Russia 2001) Find all odd positive integers $n > 1$ such that if a and b are relatively prime positive divisors of n , then $a + b - 1$ divides n .

Proof. Let p be the smallest prime divisor of n and let $n = p^k m$ with $k \geq 1$ and m relatively prime to p . By hypothesis $p + m - 1 \mid n$. Note that

$$\gcd(p + m - 1, m) = \gcd(p - 1, m) \mid \gcd(p - 1, n) = 1,$$

the last equality being a consequence of the fact that all prime factors of $p - 1$ are less than p and so they cannot divide n . Thus $p + m - 1 \mid p^k$ and $p + m - 1 = p^l$ for some $l \leq k$.

Suppose that $k \geq 2$, then $p^2 + m - 1 \mid n$ and similarly

$$\gcd(p^2 + m - 1, m) = \gcd(p^2 - 1, m) \mid \gcd(p^2 - 1, n) = 1.$$

Note that the last equality crucially uses the hypothesis that n is odd, to ensure that all prime factors of $p + 1$ are less than or equal to $\frac{p+1}{2} < p$ (since p is odd). As above, we deduce that $p^2 + m - 1 = p^j$ for some $j \leq k$. Then

$$m - 1 = p^j - p^2 = p^l - p,$$

that is $p^j + p = p^l + p^2$ or $p^l(p^{j-l} - 1) = p(p - 1)$. This immediately implies $l = 1$ and then $m = 1$. In other words, if $k \geq 2$, then n is a power of an odd prime, and it is clear that any such number is a solution of the problem.

Assume now that $k = 1$, thus necessarily $l = 1$ (as $l \leq k$ and clearly $l > 0$) and then again $m = 1$ and n is a power of p . Thus the solutions of the problem are the odd prime powers. \square

The polynomial $X^2 + X + 41$, discovered by Euler and Lagrange in the late 18th century, takes prime values for $X = 0, 1, \dots, 39$. The next example shows that it suffices to check this only for $X = 0, 1, 2, 3$.

Example 4.33. (IMO 1987) Let n be an integer greater than 2 such that $k^2 + k + n$ is a prime number for all $0 \leq k \leq \sqrt{\frac{n}{3}}$. Prove that $k^2 + k + n$ is a prime for all $0 \leq k \leq n - 2$.

Proof. Let $f(X) = X^2 + X + n$ and let p be the smallest prime factor of any of the numbers $f(0), f(1), f(2), \dots, f(n - 2)$. Suppose that the required result fails, so there is some $k \leq n - 2$ such that $f(k)$ is composite. The smallest prime factor q of $f(k)$ satisfies $q^2 \leq f(k) \leq (n - 2)^2 + n - 2 + n < n^2$, hence $q < n$. Since $p \leq q$, it follows that $p < n$.

Now let $k \in \{0, \dots, n - 2\}$ be such that $p \mid f(k)$. Let s be the remainder of k when divided by p and let $r = \min(s, p - 1 - s)$. Note that p also divides $f(s)$ and $f(p - 1 - s)$, so $p \mid f(r)$. Moreover, $r \leq \frac{p-1}{2}$, thus

$$f(r) \leq n + \left(\frac{p-1}{2}\right)^2 + \frac{p-1}{2} = n + \frac{p^2 - 1}{4}.$$

Since $p < n$ (as shows the first paragraph), we have $p \neq f(r)$, hence we can choose a prime factor q of $\frac{f(r)}{p}$. By minimality of p , we have $q \geq p$, hence $p^2 \leq f(r) \leq n + \frac{p^2-1}{4}$. It follows that $p < 2\sqrt{\frac{n}{3}}$ and then $r < \frac{p}{2} < \sqrt{\frac{n}{3}}$. But by assumption $f(r)$ is a prime number, contradicting the fact that it is a multiple of pq . \square

Remark 4.34. a) In 1952 Heegner proved that 41 is the largest integer A with the property that $n^2 + n + A$ is a prime for all $n = 0, 1, \dots, A - 2$ (Heilbronn proved that there are only finitely many such A 's in 1934).

b) The polynomial $36X^2 - 810X + 2753$ gives (by taking $X = 0, 1, \dots, 44$ and considering absolute values in case a number is negative) a string of 45 different prime values. Also for

$$f(X) = X^5 - 133X^4 + 6729X^3 - 158379X^2 + 1720294X - 6823316$$

the number $\frac{1}{4}|f(n)|$ is a prime for $0 \leq n \leq 56$.

Similarly $|3n^3 - 183n^2 + 3318n - 18757|$ is a prime for $0 \leq n \leq 46$.

4.2.3 Combinatorial number theory

Finally, we discuss some problems with a more combinatorial flavor. Most of these problems are fairly tricky.

Example 4.35. (Tuymaada 2005) The positive integers $1, 2, \dots, 121$ are arranged in the squares of a 11×11 table. Dima found the product of numbers in each row and Sasha found the product of the numbers in each column. Could they get the same set of 11 numbers?

Proof. The answer is negative. Consider the 12 primes

$$61, 67, 71, 73, 79, 83, 89, 97, 101, 103, 107, 109.$$

The only multiple of such a prime belonging to the set $\{1, 2, \dots, 121\}$ is the prime itself. Two of these primes, say p, q , must be in the same row. If Dima and Sasha found the same numbers, then there would be a column whose product of elements is a multiple of p, q . But then p, q would have to be in that column. Thus p, q belong to the same row and column, contradiction. \square

We will use several times the observation that if $a \mid b$, then the prime factors of ab are exactly the prime factors of b .

Example 4.36. (Kvant) Consider an infinite arithmetic progression of positive integers. Prove that there are infinitely many terms in this progression such that any two of them have the same set of prime divisors.

Proof. Say the general term of this progression is $a + nd$, with $n \geq 0$. All numbers $a(1 + d)^n$ with $n \geq 0$ are terms of this progression and they satisfy the desired condition. \square

Example 4.37. (Iran 2004) Let n be an integer greater than 1. Prove that there are n positive integers $a_1 < a_2 < \dots < a_n$ such that for all nonempty subsets I, J of $\{1, 2, \dots, n\}$ the numbers $\sum_{i \in I} a_i$ and $\sum_{j \in J} a_j$ have the same prime factors.

Proof. Let $a_i = i \cdot N!$ for $1 \leq i \leq n$, where N is a large integer to be chosen later. If $I \subset \{1, 2, \dots, n\}$ is a nonempty subset, then

$$\sum_{i \in I} a_i = N! \cdot \sum_{i \in I} i$$

and

$$1 \leq \sum_{i \in I} i \leq 1 + 2 + \dots + n = \frac{n(n+1)}{2}.$$

Thus if we choose $N = \frac{n(n+1)}{2}$, then the prime factors of $\sum_{i \in I} a_i$ are exactly the primes dividing $N!$, and this is independent of the choice of I . \square

Example 4.38. Let p be a prime, let $r \in \{1, 2, \dots, p-1\}$ and let $a_1, a_2, \dots, a_r \in \{1, 2, \dots, p-1\}$. Consider the remainders of all numbers $\sum_{i \in S} a_i$ when divided by p , over all subsets S of $\{1, 2, \dots, r\}$ (including the empty set, for which the corresponding sum is 0). Prove that there are at least $r+1$ different remainders among them.

Proof. The result is clear for $r = 1$. Assume that it holds for $r = k$ and that it is not true for $r = k+1$, with $k+1 < p$. By assumption we can find pairwise distinct numbers $c_1, \dots, c_k \in \{1, 2, \dots, p-1\}$ such that $0, c_1, \dots, c_k$ all appear as remainders of some of the numbers $\sum_{i \in S} a_i$ with $S \subset \{1, 2, \dots, k\}$. Since the sums

$\sum_{i \in S} a_i$ with $S \subset \{1, 2, \dots, k+1\}$ give at most $k+1$ distinct remainders, and they contain all sums $\sum_{i \in S} a_i$ with $S \subset \{1, 2, \dots, k\}$, it follows that $0, c_1, \dots, c_k$ are all possible remainders of $\sum_{i \in S} a_i$ with $S \subset \{1, 2, \dots, k+1\}$. In particular the remainders of $a_{k+1}, a_{k+1} + c_1, \dots, a_{k+1} + c_k$ (which are pairwise different) are among $0, c_1, \dots, c_k$, and so they must be a permutation of $0, c_1, \dots, c_k$. This implies that the remainders of all numbers $0, a_{k+1}, 2a_{k+1}, \dots, (p-1)a_{k+1}$ are among $0, c_1, \dots, c_k$, which forces $p \leq k+1$, contradicting the fact that $k+1 < p$. Hence if the assertion holds for $r = k$ and $k+1 < p$, then it also holds for $k+1$. This yields the desired result. \square

Example 4.39. (Erdős-Ginzburg-Ziv theorem) Let $n > 1$ be an integer. Prove that among any $2n - 1$ integers we can choose n whose arithmetic mean is an integer.

Proof. The proof is done in two steps: we prove the theorem when n is a prime using the result established in the previous example, then we deduce the general case by an elementary argument.

Assume first that $n = p$ is a prime. We may assume that our integers $a_1, a_2, \dots, a_{2p-1}$ are between 0 and $p-1$ (by replacing them with their remainders when divided by p), and we may also assume that $a_1 \leq a_2 \leq \dots \leq a_{2p-1}$. If there is $j \in \{1, 2, \dots, p-1\}$ such that $a_{p+j} = a_{j+1}$, then we must have $a_{j+1} = a_{j+2} = \dots = a_{j+p}$ and $a_{j+1} + \dots + a_{j+p}$ is a multiple of p . So assume that $a_{j+1} \neq a_{j+p}$ for $1 \leq j < p$. By the previous example, the sums of the numbers $a_{j+p} - a_{j+1}$ (for $1 \leq j < p$) give at least p distinct remainders modulo p , i.e. they cover all possible remainders. In particular, if r is the remainder of $a_1 + a_2 + \dots + a_p$, then there is a sum giving remainder $p - r$. That is, we can find some indices $1 \leq j_1 < \dots < j_k \leq p-1$ such that $a_{p+j_1} + \dots + a_{p+j_k} - a_{j_1} - \dots - a_{j_k} + a_1 + \dots + a_p$ is a multiple of p . But this last sum is clearly equal to the sum of p numbers among a_1, \dots, a_{2p-1} . The result follows.

We now treat the general case. Since any $n > 1$ is a product of primes, it suffices to prove that if $a, b > 1$ and the result holds for a and b , then it holds for ab . Consider $2ab - 1$ integers. We choose $2a - 1$ of them, and among the chosen ones we choose a whose arithmetic mean m_1 is an integer. We now consider the remaining $2ab - 1 - a$ numbers and repeat the previous

procedure: we select $2a - 1$ such numbers (if possible), then among them we select a whose arithmetic mean m_2 is an integer, and we keep doing this $2b - 1$ times, obtaining arithmetic means m_1, \dots, m_{2b-1} of collections of a numbers. Since the result holds for b , there are b numbers among m_1, \dots, m_{2b-1} whose arithmetic mean is an integer. These b arithmetic means correspond to ab integers among the $2ab - 1$ original ones, whose arithmetic mean is an integer, proving the result for ab . \square

Example 4.40. (adapted from Iran TST 2008) Let $(a_n)_{n \geq 1}$ be a sequence of positive integers such that for all $m, n \geq 1$, all prime factors of $a_m + a_n$ are among the prime factors of $m + n$. Prove that $a_n = n$ for all n .

Proof. The key observation is that $p \mid a_n + a_m$ whenever p is a prime and m, n are positive integers such that $m + n$ is a power of p . Indeed, $a_n + a_m \geq 2$, so there is a prime $q \mid a_n + a_m$, but by hypothesis $q \mid n + m$ and so $q = p$.

We first prove that if $m \neq n$, then $a_m \neq a_n$. Suppose that $a_n = a_m$ and choose a large prime p . By the first paragraph, $p \mid a_n + a_{p-n}$, thus $p \mid a_m + a_{p-n}$ and then $p \mid m + p - n$. Thus $p \mid m - n$ for all large primes, contradiction.

Next, we prove that $|a_n - a_{n+1}| = 1$ for all $n \geq 1$. Suppose that there is n for which this is not the case and let p be a prime factor of $a_{n+1} - a_n$. Let k be such that $p^k > n$. Then by the first paragraph $p \mid a_n + a_{p^k-n}$ and so $p \mid a_n + a_{p^k-n} + a_{n+1} - a_n$. That is, $p \mid a_{p^k-n} + a_{n+1}$ and then $p \mid p^k - n + n + 1 = p^k + 1$. Since this is absurd, the claim is proved.

Finally, the previous two paragraphs yield $a_{n+2} - a_{n+1} = a_{n+1} - a_n = c$ for all n and some $c \in \{-1, 1\}$. Since the sequence takes positive values, we must have $c = 1$ and so $a_n = n + k$ for some $k \geq 0$. Take a large prime p . Choose positive integers m, n such that $p \mid m + n + 2k$. Then $p \mid a_m + a_n$ and so $p \mid m + n$. Subtracting, we deduce that $p \mid 2k$ for all large primes p , thus $k = 0$ and we are done. \square

Remark 4.41. The original problem was weaker, the hypothesis being $a_m + a_n \mid (m + n)^k$ for all $m, n \geq 1$, where k is fixed.

Example 4.42. (IMO Shortlist 2007) Find all sequences $(a_n)_{n \geq 1}$ of positive integers such that:

- a) Each positive integer appears at least once in the sequence a_1, a_2, \dots ;

b) $a_n + a_m$ and a_{n+m} have the same set of prime divisors for all $n, m \geq 1$.

Proof. We will prove in a sequence of steps that only the sequence $a_n = n$ is a solution of the problem. Note that if a prime p divides $a_{n_1}, a_{n_2}, \dots, a_{n_k}$, then it also divides $a_{n_1+n_2}, a_{n_3}, \dots, a_{n_k}$ and continuing like this we obtain that $p \mid a_{n_1+n_2+\dots+n_k}$.

First, we prove that $a_1 = 1$. Otherwise, there is a prime p dividing a_1 . Using the observation in the first paragraph, we obtain $p \mid a_n$ for all n , contradicting condition a). Hence $a_1 = 1$.

Next, we prove that $\gcd(a_n, a_{n+1}) = 1$ for all n . Suppose that this is not the case and let $n \geq 1$ and let p be a prime dividing a_n and a_{n+1} . Then p also divides $a_{xn+y(n+1)}$ for all $x, y \geq 0$ (see the first paragraph). Now all integers $m \geq n(n+1)$ can be written in the form $xn + y(n+1)$ (see example 3.42), so we obtain that all but finitely many terms of the sequence a_1, a_2, \dots are multiples of p , contradicting again assumption a).

We are now ready to prove that $|a_n - a_{n+1}| = 1$ for all n . Suppose that there is n for which this is not the case, and choose a prime p dividing $a_n - a_{n+1}$. By hypothesis a), we can choose $m \geq 1$ such that $p \mid a_n + a_m$. Then $p \mid a_{n+m}$ and also $p \mid a_m + a_{n+1}$, thus $p \mid a_{n+m+1}$. This contradicts the fact that $\gcd(a_{n+m}, a_{n+m+1}) = 1$ (previous paragraph).

The previous paragraph shows that $a_{n+1} \in \{a_n - 1, a_n + 1\}$ for all n . Since $a_n > 0$ for all n and $a_1 = 1$, we must have $a_2 = a_1 + 1 = 2$. We are now in good shape to prove that $a_n = n$ for all n . Indeed, assume that there is n such that $a_{n+1} = a_n - 1$. Since $a_{n+1} = a_n - 1$ and $a_n + a_1 = a_n + 1$ have the same prime factors, both $a_n - 1$ and $a_n + 1$ must be powers of 2 (as $\gcd(a_n - 1, a_n + 1)$ divides 2) and so necessarily $a_n = 3$ and $a_{n+1} = 2$. Repeating the argument then yields $a_{n+2} \neq a_{n+1} - 1$, so $a_{n+2} = a_{n+1} + 1 = 3$. But then $a_n + a_2 = 5$ and $a_{n+2} = 3$ don't have the same prime factors, a contradiction. Hence $a_{n+1} = a_n + 1$ for all n and finally $a_n = n$ for all n . \square

Example 4.43. (Kvant M 1863) Consider the sequence of positive integers $(a_n)_{n \geq 1}$ such that $a_1 = 1, a_2 = 2$ and for all $n \geq 3$ the number a_n is the least positive integer different from a_1, a_2, \dots, a_{n-1} which is not relatively prime with a_{n-1} . Prove that this sequence contains all positive integers.

Proof. The solution is based on the following two lemmas:

Lemma 4.44. *The sequence $(a_n)_{n \geq 1}$ contains infinitely many even integers.*

Proof. Assume the contrary, i.e there exists m such that $a_m, a_{m+1}, a_{m+2}, \dots$ are all odd. Since all terms of the sequence are different there exists $k \geq m$ such that $a_k < a_{k+1}$ and $a_1, a_2, \dots, a_{m-1} < a_k$. Let p be the least prime divisor of a_k . Then $a_{k+1} \geq a_k + p$ since otherwise a_k and a_{k+1} are relatively prime. But $a_k + p$ is an even integer and therefore $a_{k+1} > a_k + p$. This shows that a_{k+1} is not the least positive integer different from a_1, a_2, \dots, a_k which is not relatively prime with a_k , a contradiction. \square

Lemma 4.45. *If the sequence $(a_n)_{n \geq 1}$ has infinitely many terms divisible by a prime p then it contains all positive integers divisible by p .*

Proof. Let k be a positive integer such that pk is not a term of the given sequence. Let m be such that $a_n > pk$ for all $n \geq m$. There is a term a_s of the sequence divisible by p with $s > m$. By the definition of a_{s+1} it follows that $a_{s+1} \leq pk$, a contradiction. \square

Now we turn to the solution of the problem. By lemmas 4.44 and 4.45 the sequence contains all positive even integers. Then by 4.45 it follows that for every prime p the sequence contains all positive integers divisible by p . This solves the problem. \square

4.3 Infinitude of primes

Having seen several examples of prime and composite numbers, we will focus on the problem of proving that there are infinitely many primes. One obvious approach to this problem is to find explicitly a sequence that contains infinitely many primes. This sounds easy, but it is not: for many sequences that naturally appear in number theory it is not known whether they contain infinitely many prime numbers.

4.3.1 Looking for primes in classical sequences

One of the most natural sequences are polynomial ones, so let us start with them. Let $f(X) = a_0 + a_1X + \dots + a_nX^n$ be a nonconstant polynomial with integer coefficients and with positive leading coefficient a_n , so that $f(k)$ is a positive integer for k large enough. We would like to know whether the sequence $f(1), f(2), \dots$ contains infinitely many primes. There is an obvious obstruction for this to happen: if there is an integer $d > 1$ dividing all the numbers $f(1), f(2), \dots$, then there are only finitely many primes in the previous sequence. Also, if we can write f as a product of two nonconstant polynomials with integer coefficients, then again there can be only finitely many primes in the sequence $f(1), f(2), \dots$. A remarkable and wide open conjecture is that these two obstructions are the only ones:

Conjecture 4.46. *Let f be a nonconstant polynomial with integer coefficients and positive leading coefficient. Assume that:*

- a) there is no integer $d > 1$ dividing all the numbers $f(1), f(2), \dots$*
- b) f is not the product of two nonconstant polynomials with integer coefficients.*

Then $f(n)$ is a prime number for infinitely many positive integers n .

To give a hint on how difficult the previous conjecture is, let us mention that there is **not a single** polynomial of degree greater than 1 for which the conjecture is proved! There are also versions of the previous conjecture, which involve several polynomials f_1, \dots, f_k and ask that $f_1(n), \dots, f_k(n)$ should be simultaneously prime for infinitely many n . A famous such conjecture is

Conjecture 4.47. *(Hardy-Littlewood prime k -tuple conjecture) Let $a_1, \dots, a_k, b_1, \dots, b_k$ be integers such that $\gcd(a_j, b_j) = 1$ for $1 \leq j \leq k$ and such that for any prime $p \leq k$ there is $x \in \mathbf{Z}$ such that p does not divide any of the numbers $a_1x + b_1, \dots, a_kx + b_k$. Then there are infinitely many n for which $a_1n + b_1, \dots, a_kn + b_k$ are all prime numbers.*

Remark 4.48. Granville proved the following amazing result: if the previous conjecture holds, then there are infinite sets A, B of positive integers such that for all $a \in A$ and $b \in B$ the number $a + b$ is prime! Also, he proved that the previous conjecture implies the existence of an infinite set A such that for all

$a, b \in A$ the number $\frac{a+b}{2}$ is prime. It is known (this is a deep theorem of Balog) that for any n there is a set A of n primes such that for all $a, b \in A$ the number $\frac{a+b}{2}$ is prime, and all these prime numbers are pairwise distinct.

Even the case when f has degree 1 in conjecture 4.46 is highly nontrivial: in this case the conjecture was proved by Dirichlet. Let us restate his remarkable and very deep result:

Theorem 4.49. (*Dirichlet's theorem*) Let a, b be relatively prime integers with $a > 0$. The arithmetic progression $(an+b)_{n \geq 0}$ contains infinitely many primes.

One can also consider the problem of understanding the arithmetic progressions all of whose terms are primes. It is an easy exercise left to the reader to check that there cannot be such an infinite arithmetic progression. On the other hand one can produce arithmetic progressions of fairly large length consisting exclusively of primes: the smallest 10-term arithmetic progression consisting of primes is $199 + 210n$ for $0 \leq n \leq 9$, the smallest 21-term arithmetic progression of primes is $5749146449311 + 26004868890n$ for $0 \leq n \leq 20$, and an arithmetic progression of primes with 26 terms is

$$43142746595714191 + 5283234035979900n$$

with $0 \leq n \leq 25$. One can easily see that they involve huge numbers for the common difference (and also the first term). The next example explains this partially:

Example 4.50. (Thébault's theorem) An increasing arithmetic progression of length $n > 2$ consists of prime numbers. Prove that the common difference is a multiple of the product of all primes less than n .

Proof. Suppose that a, d are positive integers such that $a, a+d, \dots, a+(n-1)d$ are primes. We want to prove that any prime $p < n$ divides d . Assume that $p < n$ does not divide d . Note that $\gcd(a, d) = 1$, otherwise $\gcd(a, d) > 1$ would divide both $a, a+d$ and so $a = a+d = \gcd(a, d)$, a contradiction. Since p does not divide n , the numbers $a, a+d, \dots, a+(p-1)d$ give pairwise distinct remainders when divided by p , so one of the remainders must be 0 and $a+jd$ is divisible by p for some $j \in \{0, 1, \dots, p-1\}$. Since $p \leq n$, $a+jd$ is

a prime and so necessarily $p = a + jd \geq a$ and $a < n$ (since $p < n$). But then $a + ad = a(1 + d)$ is a prime and so $a = 1$, a contradiction with the fact that a is a prime. \square

Here is a nice application of the result established in the previous example:

Example 4.51. (Tournament of the Towns 2007) Find all increasing arithmetic progressions consisting only of prime numbers, such that the number of terms is larger than the common difference.

Proof. Let $a, a + d, \dots, a + (n - 1)d$ be an arithmetic progression as in the statement of the problem, so $n > d$. Let $(p_n)_{n \geq 1}$ be the increasing sequence of primes and let k be such that $p_k < n \leq p_{k+1}$. By Thébault's theorem $p_1 \dots p_k$ divides d and so

$$p_1 \dots p_k \leq d < n \leq p_{k+1}.$$

If $p_1 \dots p_k > 2$, then $p_1 \dots p_k - 1 \geq p_{k+1}$ since $p_1 \dots p_k - 1$ must have a prime factor, and this prime factor cannot be p_1, \dots, p_k . So if $p_1 \dots p_k > 2$, then $p_1 \dots p_k > p_{k+1}$, contradicting our assumption. It follows that $p_1 \dots p_k = 2$, then $k = 1$ and $n \leq 3$. Hence we must have $n = 3$, $d < n = 3$ and $a, a + d, a + 2d$ are all primes. If $d = 1$, this is impossible since $a, a + 1$ being primes forces $a = 2$, but then $a + 2$ is not a prime. If $d = 2$, then $a, a + 2, a + 4$ should be primes. One of these numbers is a multiple of 3 and this immediately implies that $a = 3$. Hence the problem has a unique solution, the progression 3, 5, 7. \square

The following amazing (and very deep) theorem was proved in 2004, solving a problem that was open for at least 200 years:

Theorem 4.52. (Green-Tao) *For any $n \geq 3$ there is an arithmetic progression of length n consisting of prime numbers.*

Remark 4.53. Just to see how powerful this theorem is, let us mention a few straightforward consequences which would be extremely hard to prove otherwise...

a) For any n there is a set A of n primes such that for all $a, b \in A$ the number $\frac{a+b}{2}$ is prime, and all these prime numbers are pairwise distinct. Here is such a set for $n = 12$:

$A = \{71, 1163, 1283, 2663, 4523, 5651, 9311, 13883, 13931, 14423, 25943, 27611\}$.

As we have already mentioned, this is a theorem of Balog, proved before the Green-Tao theorem. Using the Green-Tao theorem, this becomes a simple exercise: consider an arithmetic progression $a + jd$ of primes for $0 \leq j \leq 2^{n+1}$, and let A be the set of numbers $a + (2^j - 1)d$ for $1 \leq j \leq n$.

b) It follows from Green-Tao that for any k and d one can find a polynomial f with integer coefficients of degree d such that $f(0), f(1), \dots, f(k)$ are all primes. Indeed, if $a + jd$ are primes for $0 \leq j \leq k^d$, the polynomial $bX^d + a$ works. It is harder to solve the similar problem with monic polynomials.

c) Yet another consequence of the Green-Tao theorem: there are arbitrarily large sets of integers A such that the average of the elements of any nontrivial subset of A is a prime. Moreover, we can impose that these primes are pairwise distinct.

Indeed, first without the restriction of the primes distinct, the construction is easy: take an arithmetic progression of primes $a + jd$, $0 \leq j < k := n \cdot n!$ and set

$$A = \{a + jn!d \mid 0 \leq j < n\}.$$

The average of the numbers $a + jn!d$ for $j \in S \subset \{1, 2, \dots, n\}$ is $a + d(\sum_{x \in S} x) \frac{n!}{|S|}$ and this is a number belonging to our arithmetic progression of primes, so it is a prime: indeed

$$0 \leq n! \frac{\sum_{x \in S} x}{|S|} < n \cdot n! = k.$$

If we want the primes to be pairwise distinct we employ the following trick. Consider a set $B = \{b_1 < \dots < b_n\}$ such that all averages of all subsets of B are pairwise distinct (for instance take $b_i = (i + 1)!$ for $1 \leq i \leq n$), then take $k = (b_n - b_1)n!$, an arithmetic progression of primes $a + jd$ as above and set $A = \{a + (b_j - b_1)n!d \mid 1 \leq j \leq n\}$. For instance, for $n = 4$ we have the set 5, 17, 89, 1277, for $n = 5$ we can take the set

$$209173, 322573, 536773, 1217893, 2484733.$$

Already for $n = 7$ it is very difficult to write down an example of such a set!

Other important sequences that arise very often in arithmetic are sequences of the form $a^n + 1$ and $a^n - 1$, where $a > 1$ is a fixed integer. One may wonder

when $a^n + 1$, respectively $a^n - 1$ are primes, where for simplicity $n > 1$. Again, there are a few easy obstructions for this to happen.

Assume first that $a^n - 1$ is a prime and that n is composite, say $n = mk$ with $m, k > 1$. Then $a^m - 1 \mid a^n - 1$ and $1 < a^m - 1 < a^n - 1$, thus $a^n - 1$ is not a prime. Therefore if $a^n - 1$ is a prime, then n is a prime. Moreover, $a - 1 \mid a^n - 1$ and $a - 1 < a^n - 1$, thus necessarily $a - 1 = 1$ and $a = 2$. In other words, the only possible primes of the form $a^n - 1$ with $a, n > 1$ are those of the form $2^p - 1$ with p a prime. However, it is not true that all these numbers are primes: one can check that $23 \mid 2^{11} - 1$ and $47 \mid 2^{23} - 1$. Prime numbers of the form $2^p - 1$ are known as Mersenne primes. It is not known whether there are infinitely many such primes, and it is not even known whether the sequence $(2^p - 1)_p$, where p runs over the prime numbers, contains infinitely many composite numbers (one can prove that this is the case if there are infinitely many primes $p \equiv 3 \pmod{4}$ such that $2p + 1$ is also a prime, by proving that for such p the number $2p + 1$ divides $2^p - 1$). The largest Mersenne prime known in 2015 is $2^{74207281} - 1$, and 49 Mersenne primes are known up to now!

Assume now that $a^n + 1$ is a prime, with $a, n > 1$. If n has a proper odd divisor m , then $a^{\frac{n}{m}} + 1 \mid a^n + 1$ and $a^n + 1$ is not a prime, contradiction. Thus n is necessarily a power of 2. One very important case is when $a = 2$, then we see that the only primes of the form $2^n + 1$ are among the Fermat numbers $F_n = 2^{2^n} + 1$. Here, the situation is much worse: again, it is not known whether the sequence F_0, F_1, \dots contains infinitely many primes or infinitely many composite numbers, and we only know 5 primes in this sequence: F_0, F_1, F_2, F_3, F_4 . This is in stark contrast with Fermat's original conjecture that all Fermat numbers are primes, a conjecture which was disproved by Euler, who proved that $641 \mid F_5 = 2^{32} + 1$ (actually $F_5 = 641 \cdot 6700417$; see also example 2.12). The only Fermat numbers whose prime factorization is known are F_0, F_1, \dots, F_{11} (even though one knows that F_n is composite for $5 \leq n \leq 32$, no prime factor of F_{20} or F_{24} is known!).

Yet another sequence which appears very often in number theory is $(n! + 1)_{n \geq 1}$. Again, it is not known whether this sequence contains infinitely many primes, even though one knows that it contains infinitely many composite numbers (this would be hard to prove at this moment, but we will see later on that

$n + 1 \mid n! + 1$ when $n + 1$ is a prime, a result known as Wilson's theorem, and this immediately implies the desired result). We will however use this sequence below to prove that there are infinitely many primes.

4.3.2 Euclid's argument

We can summarize the discussion in the previous section by saying that many of the natural sequences appearing in arithmetic are expected to contain infinitely many primes, but we are far from being able to prove such a statement. Instead of dealing with such difficult (and most of them wide open!) problems, we present in this section Euclid's wonderful indirect argument proving there are infinitely many primes, some of the consequences of the result and some related results that can be obtained with similar (but more technically involved) arguments.

Theorem 4.54. (*Euclid*) *There are infinitely many primes.*

Proof. Note that 2 is a prime, so there is at least one prime. Assume that there are only finitely many primes, call them p_1, \dots, p_k , and consider the number $1 + p_1 \cdot \dots \cdot p_k$. It is greater than 1, so it is a product of primes. Choose one of these primes and call it q . Then $q \in \{p_1, \dots, p_k\}$, since by assumption p_1, \dots, p_k exhaust all primes. In particular, q divides $p_1 \cdot \dots \cdot p_k$. But q also divides $p_1 \cdot \dots \cdot p_k + 1$, hence q divides 1, a contradiction with the fact that $q > 1$. The result follows. \square

Remark 4.55. a) Start with $a_1 = 2$ and define a_{n+1} to be the largest prime divisor of $1 + a_1 a_2 \dots a_n$. This sequence is not monotonic since $a_{10} < a_9$. It is not known whether this sequence contains all sufficiently large primes.

b) Consider the sequence whose n th term is $1 + p_1 \cdot p_2 \cdot \dots \cdot p_n$, where $p_1 < p_2 < \dots$ is the increasing sequence of primes. The first 5 terms of this sequence are all primes: 3, 7, 31, 211, 2311. However, the 6th term $1 + 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13$ is composite (a multiple of 59). It is not known if this sequence contains infinitely many prime numbers, or if it contains infinitely many composite numbers.

There are many other ways of proving that there are infinitely many primes, based on theorem 4.13 (which ensures that every integer greater than 1 has a prime factor). For instance, suppose that $(x_n)_{n \geq 1}$ is a sequence of integers

greater than 1 and pairwise relatively prime. Let p_n be a prime divisor of x_n . Then p_1, p_2, \dots are pairwise distinct primes, thus there are infinitely many primes. We know from the previous section how to construct many such sequences $(x_n)_{n \geq 1}$: for instance $x_n = 2^{2^n} + 1$ the n th Fermat number, or Sylvester's sequence defined by $x_1 = 2$ and $x_{n+1} = x_n^2 - x_n + 1$, etc.

The next examples either imitate or refine Euclid's argument. Most of them are crucially dependent on the uniqueness of the prime factorization of an integer.

Example 4.56. Let $n > 2$ be an integer. Prove that there are infinitely many primes p such that n does not divide $p - 1$. In particular (by taking $n = 3$ and $n = 4$), there are infinitely many primes of the form $3k + 2$ and infinitely many primes of the form $4k + 3$.

Proof. We will imitate Euclid's argument. Note that 2 is such a prime. Next, assume that p_1, \dots, p_k are all primes p for which n does not divide $p - 1$. Then $N = np_1 \dots p_k - 1$ is an integer greater than 1 and so it is a product of primes $N = q_1 \dots q_r$. Since N is relatively prime to p_1, \dots, p_k , none of q_1, \dots, q_r is equal to one of the numbers p_1, \dots, p_k , thus we must have $q_i \equiv 1 \pmod{n}$ for $1 \leq i \leq r$. But then

$$N = q_1 \dots q_r \equiv 1 \pmod{n},$$

and since $N \equiv -1 \pmod{n}$, we obtain that $n \mid 2$, a contradiction. \square

Example 4.57. (Romania TST 2003) Let \mathcal{P} be the set of all primes, and let M be a subset of \mathcal{P} , having at least three elements. Suppose that for any proper subset A of M , all prime factors of $-1 + \prod_{p \in A} p$ belong to M . Prove that $M = \mathcal{P}$.

Proof. Taking an odd element p of M and considering $p - 1$, we see that $2 \in M$. We will prove below that M is infinite, so let us take this for granted for a moment and see how we can finish the proof. Suppose that there is an odd prime $p \notin M$. Let p_1, p_2, \dots be the increasing sequence of elements of M and consider the numbers $p_1 - 1, p_1 p_2 - 1, p_1 p_2 p_3 - 1, \dots$. Two of them must give the same remainder when divided by p , and so we can find $i < j$ such that p divides $p_1 \dots p_i - p_1 \dots p_j$. Since p is not p_1, \dots, p_i (as $p \notin M$), p must divide

$p_{i+1} \dots p_j - 1$. But by assumption all prime factors of this last number belong to M , contradiction. Hence $M = \mathcal{P}$.

Let us prove now that M is infinite. Assume the contrary and let p be the smallest odd element of M , and let x be the product of the elements of $M \setminus \{2, p\}$ (we are using here that M has at least three elements). All prime factors of x are greater than p and all prime factors of $x - 1$ and $2x - 1$ belong to M , by assumption. It follows that $2x - 1 = p^a$ and $x - 1 = 2^b p^c$ for some $a \geq 1$, $b, c \geq 0$. Since $x - 1$ and $2x - 1$ are relatively prime, we must have $c = 0$ and $x = 2^b + 1$, then $p^a = 2^{b+1} + 1$. If a is even, say $a = 2k$, then $(p^k - 1)(p^k + 1)$ is a power of 2, thus $p^k - 1$ and $p^k + 1$ are powers of 2 differing by 2, so $p^k = 3$ and $p = 3$. If a is odd, then $(p - 1)(1 + p + \dots + p^{a-1}) = 2^{b+1}$ is impossible, since $1 + p + \dots + p^{a-1}$ is odd and greater than 1. Thus $p = 3$ and $3 \in M$. We deduce that $2 \cdot 3 - 1 = 5$ must be in M , and $7 \mid 3 \cdot 5 - 1$ must divide $x = 2^b + 1$. This is impossible, since 7 does not divide $2^b + 1$ for any b . This contradiction shows that M must be infinite and the problem is solved. \square

Remark 4.58. A very similar problem was proposed at the USA TST in 2015: let M be a nonempty set of primes such that for any nonempty subset $N \subset M$, all prime factors of $1 + \prod_{p \in N} p$ are also in M . Prove that M is the set of all primes.

Example 4.59. Let $(a_n)_{n \geq 1}$ be a sequence of pairwise distinct positive integers. Suppose that there are positive integers k, c such that $a_n \leq cn^k$ for all $n \geq 1$. Prove that there are infinitely many primes p dividing at least one of the numbers a_1, a_2, \dots

Proof. Suppose that there are only finitely many such primes, call them p_1, p_2, \dots, p_s . Choose a large integer $N > c$ and consider the positive integers in $\{1, 2, \dots, c2^{Nk}\}$. There are at least 2^N terms of the sequence among these numbers, namely a_1, \dots, a_{2^N} . On the other hand, all these terms have prime factors among p_1, p_2, \dots, p_s , so can be written uniquely as $p_1^{x_1} \dots p_s^{x_s}$ for an s -tuple (x_1, \dots, x_s) of nonnegative integers. Since $p_1^{x_1} \dots p_s^{x_s} \leq c \cdot 2^{Nk} < 2^{N(k+1)}$ and $p_1^{x_1} \dots p_s^{x_s} \geq 2^{x_i}$ for all i , we deduce that $x_i < (k+1)N$ for all i . Thus there are at most $((k+1)N)^s$ such s -tuples and consequently at most $((k+1)N)^s$

numbers between 1 and $c \cdot 2^{Nk}$ all of whose prime factors are among p_1, \dots, p_s . Since a_1, \dots, a_{2N} are such numbers, we deduce that

$$((k+1)N)^s \geq 2^N.$$

This certainly does not hold if N is big enough: the left-hand side is smaller than a constant times N^s , but for N large enough $2^N > N^{s+1}$ using the inequality $2^N > \binom{N}{s+2}$ and the fact that $\binom{N}{s+2}$ is a polynomial expression of degree $s+2$ in N . The result follows. \square

Remark 4.60. As the proof clearly shows, it suffices to ensure that there is M such that each integer occurs at most M times in the sequence a_1, a_2, \dots . If f is a nonconstant polynomial with integer coefficients, the sequence $a_n = f(n)$ has this property, so the previous result shows that infinitely many primes divide at least one of the numbers $f(1), f(2), \dots$. This gives an alternative proof of theorem 4.67 below.

The existence of infinitely many primes is a very useful tool in constructive problems. Here are a few typical examples of problems whose statement has nothing to do with primes and whose solution crucially depends on the existence of infinitely many primes:

Example 4.61. (Tournament of the Towns 2006) For each positive integer n let b_n be the denominator of $1 + \frac{1}{2} + \dots + \frac{1}{n}$ when written in lowest terms. Prove that $b_{n+1} < b_n$ for infinitely many n .

Proof. We will prove that $n = p^2 - p - 1$ is a solution of the problem for each odd prime p . First, we claim that p does not divide b_{n+1} . Indeed, the only fractions among $\frac{1}{2}, \dots, \frac{1}{n}$ whose denominator is a multiple of p are $\frac{1}{p}, \frac{1}{2p}, \dots, \frac{1}{(p-1)p}$, but their sum is a fraction whose denominator is not a multiple of p , since $\frac{1}{p} + \frac{1}{(p-1)p} = \frac{1}{p-1}$, $\frac{1}{2p} + \frac{1}{(p-2)p} = \frac{1}{2(p-2)}$, etc.

Next, let a_n be the numerator of $1 + \frac{1}{2} + \dots + \frac{1}{n}$, so that

$$\frac{a_{n+1}}{b_{n+1}} = \frac{a_n}{b_n} + \frac{1}{p(p-1)}.$$

Thus

$$\frac{a_n}{b_n} = \frac{p(p-1)a_{n+1} - b_{n+1}}{p(p-1)b_{n+1}}.$$

If $d = \gcd(p(p-1)a_{n+1} - b_{n+1}, p(p-1)b_{n+1})$, then d divides $p^2(p-1)^2a_{n+1}$ and $p(p-1)b_{n+1}$, hence d divides $p^2(p-1)^2$. But p does not divide d , since it does not divide b_{n+1} . Hence $d \mid (p-1)^2$ and so

$$b_n \geq b_{n+1} \frac{p(p-1)}{(p-1)^2} > b_{n+1}.$$

The result follows. \square

Example 4.62. (IMO Shortlist 2011) Let $n \geq 1$ be an odd integer. Find all functions $f : \mathbf{Z} \rightarrow \mathbf{Z}$ such that $f(x) - f(y)$ divides $x^n - y^n$ for all integers x, y .

Proof. It is clear that all functions of the form $f(x) = \varepsilon x^d + c$ with $\varepsilon \in \{-1, 1\}$ and d a positive divisor of n are solutions. We will prove that these are all solutions of the problem. Note that if f is such a function, then $f + c$ has the same property for any integer c , hence we may assume that $f(0) = 0$.

If p is a prime, then $f(p) - f(0)$ divides p^n , thus $f(p) \mid p^n$ and so $f(p) = \pm p^d$ for some sign \pm and some $0 \leq d \leq n$. We deduce that there is a sign ε and some fixed $0 \leq d \leq n$ such that $f(p) = \varepsilon p^d$ for infinitely many primes, call them $p_1 < p_2 < \dots$. We may assume that $\varepsilon = 1$, by replacing f with $-f$. Now $p_1^d - p_2^d$ divides $p_1^n - p_2^n$ by hypothesis, hence d divides n (by corollary 3.36). Write $n = kd$.

We will now prove that $f(x) = x^d$ for all x . Fix an integer x . Then $f(x) - p_i^d$ divides $x^n - p_i^n$ and it also divides $f(x)^k - p_i^{dk} = f(x)^k - p_i^n$. Thus $f(x) - p_i^d$ divides $f(x)^k - x^n$, and this for all $i \geq 1$. Note that $d \neq 0$, since d divides n . Hence $d > 0$, and since $f(x) - p_i^d$ divides $f(x)^k - x^n$ for all i , it follows that $f(x)^k - x^n$ has infinitely many divisors, thus it must be 0, and then $f(x) = x^d$ (since n is odd). \square

Example 4.63. (USA TST 2010) Let P be a polynomial with integer coefficients such that $P(0) = 0$ and

$$\gcd(P(0), P(1), P(2), \dots) = 1.$$

Prove that for infinitely many n

$$\gcd(P(n) - P(0), P(n+1) - P(1), P(n+2) - P(2), \dots) = n.$$

Proof. Let us try to study first

$$d_n = \gcd(P(n) - P(0), P(n+1) - P(1), \dots)$$

for any polynomial P with integer coefficients. Let q be a prime factor of d_n , so that $P(n+k) \equiv P(k) \pmod{q}$ for all k , i.e. P is n -periodic modulo q . But P is also q -periodic modulo q . Thus, if $\gcd(q, n) = 1$, then P is 1-periodic modulo q (by Bézout's theorem) and so q divides $P(n+1) - P(n)$ for all n . Then q divides $P(n) - P(0)$ for all n , so if $P(0) = 0$, then q must divide $\gcd(P(0), P(1), \dots)$. In particular, for our polynomial we must have $q|n$ for any prime factor q of d_n .

The previous paragraph suggests taking for n a power of a prime, say $n = p^N$. Then we saw that d_n is also a power of p . Note that d_n is a multiple of n , since n divides $P(n+k) - P(k)$ for all k . It remains to see if we can have $p^{N+1} | P(k+p^N) - P(k)$ for all k . Since¹

$$P(k+p^N) \equiv P(k) + p^N P'(k) \pmod{p^{N+1}},$$

this would imply that p divides $P'(k)$ for all k . Now we see how to choose our numbers n : pick and fix once and for all a value k such that $P'(k) \neq 0$. If p is sufficiently large, then p does not divide $P'(k)$. For any such p , the previous arguments show that $d_n = n$ for all $n = p^N$. The conclusion follows. \square

Example 4.64. (Erdős) Let A be a set of n nonzero integers. Prove that A contains a subset B with more than $\frac{n}{3}$ elements, such that the sum of any two elements of B (not necessarily distinct) is not an element of B .

Proof. Let the elements of A be a_1, a_2, \dots, a_n and let $p = 3k + 2$ be a prime number greater than $\max |a_i|$ (such a prime exists thanks to example 4.56). For any $i \in \{1, \dots, n\}$ the numbers $a_i, 2a_i, \dots, pa_i$ form a complete system of residues modulo p (since $a_i \neq 0$ and $|a_i| < p$). It follows that for each $1 \leq i \leq n$ one can find $k+1$ numbers among $a_i, 2a_i, \dots, na_i$ which are congruent to $k+1, k+2, \dots, 2k+1$ modulo p .

¹We recall that P' is the derivative of P .

For each $1 \leq j \leq p$ let B_j be the set of those a_i for which the remainder of ja_i modulo p belongs to $\{k+1, \dots, 2k+1\}$. It follows from the first paragraph that

$$\sum_{j=1}^p |B_j| = (k+1)n,$$

hence we can find j with $|B_j| \geq \frac{k+1}{3k+2}n > \frac{n}{3}$. It remains to check that the sum of two elements of B_j is not in B_j . Suppose that $x, y, z \in B$ satisfy $x + y = z$. By definition the remainders of jx, jy, jz when divided by p are in $\{k+1, k+2, \dots, 2k+1\}$ and one of them is the sum of the remaining two. This is clearly impossible, since no two elements of $\{k+1, \dots, 2k+1\}$ add up to a third element of this set. Thus B_j satisfies all desired conditions. \square

Example 4.65. (Iran 2011) Find all sequences $(a_n)_n$ of positive integers such that $na_n + ma_m + 2mn$ is a perfect square for all positive integers m, n .

Proof. The key is to prove that $a_p = p$ for sufficiently large primes p . Assume that this is the case for a moment, and fix a positive integer n . By assumption $na_n + 2np + p^2$ is a square for all sufficiently large primes p , hence there is a prime p_0 and a sequence of positive integers $(b_p)_{p > p_0}$ such that $na_n + 2np + p^2 = (p + b_p)^2$. Then $2pb_p + b_p^2 = 2np + na_n$ and so $b_p < n + na_n$. Hence the sequence $(b_p)_{p > p_0}$ is bounded and since p divides $b_p^2 - na_n$ for all $p > p_0$, it follows that $b_p^2 = na_n$ for p large enough. But then the relation $2pb_p + b_p^2 = 2np + na_n$ yields $b_p = n$ and finally $a_n = n$. Hence, modulo the initial claim, we find that there is a unique sequence, namely $a_n = n$ for all n .

Now, let us prove that $a_p = p$ for all large primes p . Since $2pa_p + 2p^2$ is a square and a multiple of p , it must be a multiple of p^2 , hence p divides a_p if $p > 2$.

Next, we prove that na_n is a square for all n . Indeed, fixing n and choosing $m = (na_n)^2$, we see that $na_n + ma_m + 2mn = na_n(1 + xna_n)$ for some integer x . Since na_n and $1 + xna_n$ are relatively prime positive integers and their product is a perfect square, na_n must be a perfect square.

Finally, by the previous two paragraphs we can write $pa_p = (px_p)^2$ for some positive integer x_p , and this holds for all primes $p > 2$. Since $pa_p + 2p + a_1$ is a square, it can be written as $(px_p + y_p)^2$ for some positive integer y_p . Now

$2px_p y_p + y_p^2 = 2p + a_1$. If $2p > a_1$, then necessarily $x_p y_p < 2$, hence $x_p = y_p = 1$ and so $a_p = p$ whenever $2p > a_1$ and $p > 2$. This finishes the proof. \square

Example 4.66. For any integer $n > 1$, let $P(n)$ denote the largest prime divisor of n . Prove that there are infinitely many positive integers n for which

$$P(n) < P(n+1) < P(n+2).$$

Proof. We will prove that for each prime $p > 2$ we can find $k \geq 1$ such that

$$P(p^{2^k} - 1) < P(p^{2^k}) = p < P(p^{2^k} + 1),$$

which will be enough to conclude. The numbers $(p^{2^k} + 1)_k$ are pairwise relatively prime² and not divisible by 4, so the sequence $(P(p^{2^k} + 1))_{k \geq 1}$ is unbounded. Hence there is a smallest k for which $P(p^{2^k} + 1) > p$. We will prove that $P(p^{2^k} - 1) < p$. Otherwise, there is a prime $q \geq p$ such that $q \mid p^{2^k} - 1 = (p-1)(p+1)(p^2+1)\dots(p^{2^{k-1}}+1)$. Clearly $q \neq p$ and since p is odd and $q \geq p$, q does not divide $p+1$. Thus q divides one of the numbers $p^{2^j} + 1$ with $1 \leq j < k$, and $P(p^{2^j} + 1) > p$, contradicting the minimality of k . \square

Here is yet another very short proof of the existence of infinitely many primes. Consider the number $x_n = n! + 1$. Since $x_n > 1$, x_n has at least one prime divisor, say p_n . Since p_n cannot divide $1, 2, \dots, n$ (as otherwise p_n divides both $n!$ and $n! + 1$, impossible), we must have $p_n > n$. Hence the sequence $(p_n)_{n \geq 1}$ has infinitely many distinct terms and the result follows. The proof of the following very useful result is a variation on the previous argument:

Theorem 4.67. (Schur) *Let f be a nonconstant polynomial with integer coefficients. There are infinitely many primes dividing at least one nonzero term of the sequence $f(1), f(2), f(3), \dots$*

Proof. Let $f(X) = a_0 + a_1X + \dots + a_nX^n$, with $a_n \neq 0$ and $n \geq 1$. If $a_0 = 0$, then any prime p divides $f(p)$ and $f(p) \neq 0$ for all sufficiently large p , thus the result is clear in this case. Assume that $a_0 \neq 0$ and observe that

$$f(a_0X) = a_0 + a_0a_1X + \dots + a_0^n a_n X^n = a_0(1 + a_1X + \dots + a_0^{n-1} a_n X^n).$$

²This can be proved in the same way as for Fermat numbers.

The polynomial $g(X) = 1 + a_1X + \dots + a_0^{n-1}a_nX^n$ is nonconstant, hence there is an integer k_0 such that for all $x \geq k_0$ we have $|g(x)| \geq 2$. Pick any prime p_k dividing $g(k!)$, for $k \geq k_0$. Then p_k divides $g(k!)$ and $k! \mid g(k!) - 1$, thus p_k is relatively prime to $k!$ and so $p_k > k$. Moreover, p_k divides $f(a_0k!)$ for $k \geq k_0$ and since $p_k > k$, the result follows. \square

The following examples illustrate the previous theorem.

Example 4.68. (Iran 2004) Find all polynomials f with integer coefficients such that $f(m)$ and $f(n)$ are relatively prime whenever m and n are relatively prime positive integers.

Proof. Note that the polynomials $\pm X^k$, with $k \geq 0$ are solutions of the problem. We will prove that these are the only solutions. Let f be a solution and write $f(X) = X^k g(X)$ with $k \geq 0$ and $g(0) \neq 0$. If g is constant, then clearly this constant must be ± 1 . Suppose that g is not constant, hence for infinitely many primes p the congruence $g(n) \equiv 0 \pmod{p}$ has solutions. Choose such p and n , with p relatively prime to $g(0) \neq 0$ (since $g(0) \neq 0$, this holds for all but finitely many primes p). Then p does not divide n , hence n and $n+p$ are relatively prime. But then $f(n)$ and $f(n+p)$ are relatively prime, which contradicts the fact that p divides both of them (since p divides $g(n)$, it also divides $g(n+p)$). Thus g is constant and we are done. \square

Example 4.69. (Taiwan TST 2014) Let k be a positive integer. Find all polynomials $f(X)$ with integer coefficients such that $f(n)$ divides $(n!)^k$ for all positive integers n .

Proof. Replacing f with $-f$ we may assume that the leading coefficient of f is positive. If f is constant, then since $f(1) \mid 1$ we must have $f(X) = 1$, which is a solution of the problem. Assume now that f is not constant, and write $f(X) = a_0 + a_1X + \dots + a_dX^d$ with $a_d > 0$ and $d \geq 1$. Let j be the smallest nonnegative integer such that $a_j \neq 0$, thus $a_0 = \dots = a_{j-1} = 0$ and $a_j \neq 0$. Then $f(X) = X^j g(X)$ with $g(X) = a_j + a_{j+1}X + \dots + a_dX^{d-j}$. Assume that $j < d$, so that $g(X)$ is nonconstant. By hypothesis $g(n) \mid (n!)^k$ and so for any prime $p \mid g(n)$ we have $p \mid n!$ and $p \leq n$. Since g is nonconstant, theorem 4.67 yields the existence of infinitely many primes p dividing at least one of

the numbers $g(1), g(2), \dots$. Let p be such a prime and let n be the smallest positive integer for which $p \mid g(n)$. If r is the remainder of n when divided by p , then $p \mid g(n) - g(r)$ and so $p \mid g(r)$. If $r > 0$, then since $g(r) \mid (r!)^k$ we must have $p \leq r$, impossible. Thus $r = 0$ and so $p \mid g(0)$. Thus $g(0) = a_j$ is divisible by infinitely many primes and is nonzero, which is impossible. Hence our assumption that $j < d$ was wrong and $f(X) = a_d X^d$. Since $a_d > 0$ and $f(1) \mid 1$, we obtain $a_d = 1$. Then $n^d \mid (n!)^k$ for all $n \geq 1$. Choosing $n = p$ a prime, we see that $(n!)^k = p^k \cdot m$ with m not a multiple of p , thus $p^d \mid (p!)^k$ forces $d \leq k$. Conversely, if $d \leq k$ then clearly $n^d \mid (n!)^k$ for all n . Hence the solutions of the problem are $f(X) = \pm X^d$ with $d \leq k$.

Here is a slightly different argument: assume that f is not constant and has positive leading coefficient. Thus if p is a large enough prime, we have $f(p) > 1$. Let q be a prime factor of $f(p)$ and assume that $q \neq p$. Write $p = qk + r$ with $0 < r < q$. Since $q \mid f(p) = f(qk + r)$, we have $q \mid f(r) \mid r!^k$. This is impossible, since q is a prime greater than r . Thus $q = p$ and so $p \mid f(p)$. It follows that $p \mid f(0)$ for all large enough primes, thus $f(0) = 0$. Thus we can argue like in the previous solution. \square

Example 4.70. a) (Saint Petersburg 2001) Prove that there are infinitely many positive integers n such that the largest prime divisor of $n^4 + 1$ is greater than $2n$.

b) (IMO 2008) Prove that the largest prime factor of $n^2 + 1$ is greater than $2n + \sqrt{2n}$ for infinitely many positive integers n .

Proof. a) By theorem 4.67 there are infinitely many odd primes p dividing at least one of the numbers $n^4 + 1$ with $n \geq 1$. Let p be one such prime and let n be the smallest positive integer such that $p \mid n^4 + 1$. If r is the remainder of n when divided by p , then $r < p$ and $p \mid r^4 + 1$. By minimality of n (note that $r > 0$) we have $n \leq r$, thus $n \leq p - 1$, and actually $n < p - 1$, since p does not divide $(p - 1)^4 + 1$ (as p is odd). Next, $p \mid (p - 1 - n)^4 + 1$ and again by minimality of n we have $p - 1 - n \geq n$, that is $p > 2n$. Thus to any prime p as above we associated a positive integer $n_p < \frac{p}{2}$ such that $p \mid n_p^4 + 1$. Since $n_p^4 \geq p - 1$, as p varies the numbers n_p form an unbounded sequence, and the largest prime factor of $n_p^4 + 1$ is at least $p > 2n_p$, solving the problem.

b) As above, start with an odd prime p dividing one of the numbers $n^2 + 1$, with $n \geq 1$. Let n be the smallest such positive integer. As above, we obtain $n \leq \frac{p-1}{2}$. Now write $p = 2k + 1$ and $s = k - n \geq 0$. Then $p \mid 4n^2 + 4 = (2k - 2s)^2 + 4$, hence $p \mid (2s + 1)^2 + 4$. It follows that $(2s + 1)^2 + 4 \geq p = 2k + 1$, thus $2s + 1 \geq \sqrt{2k - 3}$. Now

$$p = 2k + 1 = 2n + 2s + 1 \geq 2n + \sqrt{2k - 3}.$$

Let us prove that if p is large enough, then $\sqrt{2k - 3} > \sqrt{2n}$, which is enough to conclude. The inequality $\sqrt{2k - 3} > \sqrt{2n}$ is equivalent to $k - 2 \geq n$ or $s \geq 2$. Since $p \mid (2s + 1)^2 + 4$, it suffices to take $p > 3^2 + 4 = 13$ for the argument to work. \square

Remark 4.71. We suggest the reader to try the following problem, proposed for the USAMO in 2006: let $P(n)$ be the largest prime divisor of n (with $P(\pm 1) = 1$ and $P(0) = \infty$). Find all polynomials f with integer coefficients such that the sequence $(P(f(n^2)) - 2n)_{n \geq 1}$ is bounded above.

Example 4.72. (Romania TST 2013) Prove that infinitely many prime numbers can be written as

$$\frac{(a_1^2 + a_1 - 1)(a_2^2 + a_2 - 1) \dots (a_n^2 + a_n - 1)}{(b_1^2 + b_1 - 1)(b_2^2 + b_2 - 1) \dots (b_n^2 + b_n - 1)}$$

for some positive integers $n, a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n$.

Proof. Let p_1, p_2, \dots be the prime numbers dividing at least one of the numbers $1^2 + 1 - 1, 2^2 + 2 - 1, 3^2 + 3 - 1, \dots$. We claim that p_i has the required form for all $i \geq 1$ and we prove this by strong induction on i . For $i = 1$ we have $p_1 = 5$ (it is easy to see that 2 and 3 do not divide any $n^2 + n - 1$) and this equals $\frac{2^2 + 2 - 1}{1^2 + 1 - 1}$. Let S be the set of rational numbers of the form

$$\frac{(a_1^2 + a_1 - 1)(a_2^2 + a_2 - 1) \dots (a_n^2 + a_n - 1)}{(b_1^2 + b_1 - 1)(b_2^2 + b_2 - 1) \dots (b_n^2 + b_n - 1)}$$

for some positive integers $n, a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n$ and assume that $p_1, \dots, p_{k-1} \in S$. Let us prove that $p_k \in S$. By assumption there is $n \geq 1$

such that $p_k \mid n^2 + n - 1$. Let n be the smallest positive integer such that $p_k \mid n^2 + n - 1$. Then $n < p_k - 1$ and $m = p_k - (n + 1)$ is a positive integer such that $p_k \mid m^2 + m - 1$, thus $n < m$ and $n < \frac{p_k - 1}{2}$. We deduce that $n^2 + n - 1$ is of the form $p_k s$ for some $s < p_k$. By definition, all prime factors of s are among p_1, \dots, p_{k-1} and thus all prime factors of s are in S . Clearly S is stable under product, thus $s \in S$ and then

$$p_k = \frac{n^2 + n - 1}{1^2 + 1 - 1} \cdot \frac{1}{s} \in S. \quad \square$$

4.3.3 Euler's and Bonse's inequalities

The following remarkable inequality goes back to Euler. It immediately implies the existence of infinitely many primes, and we will see that it yields a very strong estimate for the sum of the inverses of the primes not exceeding n .

Theorem 4.73. (Euler) *Let p_1, p_2, \dots, p_k be all primes not exceeding n . Then*

$$\frac{p_1}{p_1 - 1} \cdot \frac{p_2}{p_2 - 1} \cdot \dots \cdot \frac{p_k}{p_k - 1} > 1 + \frac{1}{2} + \dots + \frac{1}{n}.$$

Proof. We have for all $N \geq 1$

$$1 + \frac{1}{p_i} + \dots + \frac{1}{p_i^N} = \frac{1 - \frac{1}{p_i^{N+1}}}{1 - \frac{1}{p_i}} < \frac{1}{1 - \frac{1}{p_i}} = \frac{p_i}{p_i - 1},$$

hence

$$\prod_{i=1}^k \frac{p_i}{p_i - 1} > \prod_{i=1}^k \left(1 + \frac{1}{p_i} + \dots + \frac{1}{p_i^N} \right).$$

Expanding the product we obtain

$$\prod_{i=1}^k \frac{p_i}{p_i - 1} > \sum_{\alpha_1, \dots, \alpha_k \in \{0, 1, \dots, N\}} \frac{1}{p_1^{\alpha_1} \dots p_k^{\alpha_k}}.$$

On the other hand, by the fundamental theorem of arithmetic, all numbers $j \in [1, n]$ can be written as a product of powers of primes. If $j = q_1^{\alpha_1} \dots q_r^{\alpha_r}$ is

the prime factorization of j , then $\max(q_i) \leq j \leq n$, hence q_1, \dots, q_r are among p_1, \dots, p_k . Moreover, $n \geq j \geq 2^{\alpha_i} > \alpha_i$ for all i . Hence if we take $N = n$ we conclude that

$$\sum_{\alpha_1, \dots, \alpha_k \in \{0, 1, \dots, N\}} \frac{1}{p_1^{\alpha_1} \dots p_k^{\alpha_k}} \geq 1 + \frac{1}{2} + \dots + \frac{1}{n}.$$

Combining this with the first inequality yields the desired result. \square

Theorem 4.74. (Euler) For all $n > 1$ we have

$$\sum_{p \leq n} \frac{1}{p} > \ln \ln n - 1,$$

the sum being taken over all primes not exceeding n . In particular,

$$\sum_p \frac{1}{p} = \infty,$$

i.e. the sum of the inverses of all primes diverges.

Proof. The inequality $x \geq \ln(1+x)$ holds for all $x \geq 0$. Using it, we obtain

$$\sum_{k=1}^n \frac{1}{k} \geq \sum_{k=1}^n \ln \left(1 + \frac{1}{k} \right) = \sum_{k=1}^n (\ln(k+1) - \ln(k)) = \ln(n+1).$$

On the other hand, letting p_1, \dots, p_k the primes not exceeding n , we have (using the inequality $x \leq e^x - 1$)

$$\prod_{i=1}^k \frac{p_i}{p_i - 1} = \prod_{i=1}^k \left(1 + \frac{1}{p_i - 1} \right) \leq \prod_{i=1}^k e^{\frac{1}{p_i - 1}} = e^{\sum_{i=1}^k \frac{1}{p_i - 1}}.$$

Combining these inequalities with the one obtained in the previous theorem we obtain

$$\sum_{i=1}^k \frac{1}{p_i - 1} > \ln(\ln(n+1)).$$

It suffices therefore to prove that

$$\sum_{i=1}^k \frac{1}{p_i - 1} - \sum_{i=1}^k \frac{1}{p_i} \leq 1$$

or equivalently

$$\sum_{i=1}^k \frac{1}{p_i(p_i - 1)} \leq 1.$$

But since $p_i \geq i + 1$ we have

$$\sum_{i=1}^k \frac{1}{p_i(p_i - 1)} \leq \sum_{i=1}^k \frac{1}{i(i + 1)} = \sum_{i=1}^k \left(\frac{1}{i} - \frac{1}{i + 1} \right) < 1,$$

as desired. □

Remark 4.75. The inequality established in the previous theorem is remarkably strong. More precisely, one can prove (with quite a lot of work) that

$$\lim_{n \rightarrow \infty} \left(\sum_{p \leq n} \frac{1}{p} - \ln \ln n \right) = 0.2614\dots,$$

so in terms of growth the inequality is essentially optimal!

It will be convenient for the next examples and results to have a notation for the n th prime.

Definition 4.76. If n is a positive integer, we let p_n be the n th prime number, thus $p_1 = 2 < p_2 = 3 < p_3 = 5 < p_4 = 7 < \dots$ is the increasing infinite sequence of primes.

Example 4.77. Prove that for all $n \geq 1$ we have

$$\sum_{k=1}^n \frac{1}{p_k^2} < \frac{49}{100}.$$

Proof. For $k \geq 4$ we have $p_k \geq 2k - 1$ hence $p_k^2 > 4k(k - 1)$ and so

$$\frac{1}{p_k^2} < \frac{1}{4} \left(\frac{1}{k-1} - \frac{1}{k} \right).$$

We deduce that

$$\sum_{k=1}^n \frac{1}{p_k^2} < \frac{1}{4} + \frac{1}{9} + \frac{1}{25} + \sum_{k=4}^n \frac{1}{4} \left(\frac{1}{k-1} - \frac{1}{k} \right) < \frac{1}{4} + \frac{1}{9} + \frac{1}{25} + \frac{1}{12}.$$

The last expression is equal to $\frac{4}{9} + \frac{4}{100}$ and so we are reduced to checking that $\frac{4}{9} < \frac{45}{100} = \frac{9}{20}$, which in turn is equivalent to $81 > 80$. \square

Remark 4.78. The phenomena appearing in the last two examples are very similar to the behavior of the sequence of positive integers: we have

$$\sum_{k=1}^n \frac{1}{k^2} < 2$$

for all $n \geq 1$, but there is no real number M such that

$$\sum_{k=1}^n \frac{1}{k} < M$$

for all n . In other words $\sum_{k \geq 1} \frac{1}{k} = \infty$.

Example 4.79. Let p_n be the n th prime. Prove that

- a) $p_n > 2n$ for $n \geq 5$.
- b) $p_n > 3n$ for $n \geq 12$.

Proof. a) We prove this by induction. For $n = 5$ we have $p_5 = 11 > 2 \cdot 5$, so assume that $p_n > 2n$ and let us prove that $p_{n+1} > 2(n+1)$. But p_{n+1} is odd and greater than p_n , which is also odd, hence $p_{n+1} \geq p_n + 2 > 2n + 2$, and the result follows.

b) Again, we prove this by induction. A direct computation shows that $p_{12} = 37 > 3 \cdot 12$, so assume that $p_n > 3n$ and let us prove that $p_{n+1} > 3(n+1)$. As before, $p_{n+1} \geq p_n + 2 \geq 3n + 1 + 2 = 3(n+1)$. Since $3(n+1)$ is not a prime, the previous inequality cannot be an equality and so $p_{n+1} > 3(n+1)$. \square

Remark 4.80. It is true, but not easy to prove that for any positive integer k there is n_k such that for all $n > n_k$ we have $p_n > kn$. We will see a proof of this result later on.

The next example uses a similar argument to Euclid's one, but it is technically more involved. We will use it to give a very elementary proof of a famous inequality of Bonse, and then give some interesting arithmetic applications of this inequality.

Example 4.81. Prove that if $n \geq 4$ then

$$p_1 p_2 \dots p_n \geq p_{p_n+n-2} + p_1 p_2 \dots p_{n-1} + p_n.$$

Proof. Write the inequality as

$$p_1 \dots p_{n-1} (p_n - 1) - p_n \geq p_{p_n+n-1}$$

and consider the numbers $x_k = kp_1 p_2 \dots p_{n-1} - p_n$ for $2 \leq k < p_n$. We need to prove that $x_{p_n-1} \geq p_{p_n+n-2}$.

First, note that $p_1 \dots p_{n-1} - 1$ is greater than 1 and is relatively prime to p_1, p_2, \dots, p_{n-1} , so all of its prime divisors are at least p_n , in particular $p_1 \dots p_{n-1} \geq p_n + 1$ and so $x_k \geq 2(p_n + 1) - p_n = p_n + 1$ for $2 \leq k < p_n$.

Next, we claim that the numbers x_k are pairwise relatively prime. Indeed, if a prime q divides x_k and x_j for some $2 \leq j < k < p_n$, then it divides $x_k - x_j = (k - j)p_1 \dots p_{n-1}$. Now $q \neq p_1, \dots, p_{n-1}$ since none of these primes divide x_k (as they don't divide p_n), so $q \mid k - j$. But then $q < p_n$ and so $q \in \{p_1, \dots, p_{n-1}\}$, a contradiction.

Now, let q_k be the smallest prime factor of x_k , then q_2, \dots, q_{p_n-1} are $p_n - 2$ pairwise distinct prime numbers by the previous paragraph, and they are all larger than p_n , since clearly x_k is relatively prime to $p_1 \dots p_n$. Thus

$$\max(q_2, \dots, q_{p_n-1}) \geq p_{n+p_n-2}$$

and thus

$$x_{p_n-1} \geq \max(q_2, \dots, q_{p_n-1}) \geq p_{n+p_n-2},$$

as desired. □

Example 4.82. (Bonse's inequality) For $n \geq 4$ we have

$$p_1 p_2 \dots p_n > p_{n+1}^2.$$

and for $n \geq 5$ we have

$$p_1 p_2 \dots p_{n-1} > p_{n+1}^2.$$

Proof. One can check the first inequality for $n = 4$ without any problem, so it suffices to prove the stronger inequality $p_1 \dots p_{n-1} > p_{n+1}^2$ for $n \geq 5$. Assume first that $n \geq 12$ and let $k = \lfloor \frac{n}{2} \rfloor$ so $2k \leq n \leq 2k + 1$. Then $k \geq 6$ and

$$p_1 \dots p_{n-1} > p_1 \dots p_{2k-2} > (p_1 \dots p_{k-1})^2.$$

Using the previous example, the last quantity is greater than $p_{p_{k-1}+k-3}^2$ and so it suffices to check that $p_{k-1} + k - 3 \geq n + 1$, or the stronger inequality $p_{k-1} + k - 3 \geq 2k + 2$. This reduces to $p_{k-1} \geq k + 5$ and is easily checked for $k \geq 6$. Hence the result is proved for $n \geq 12$. Next, Note that

$$p_{13}^2 = 41^2 < 2000 < 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 = p_1 p_2 p_3 p_4 p_5$$

hence the result holds for $5 \leq n \leq 12$ too. □

We illustrate the usefulness of Bonse's inequality with two examples:

Example 4.83. a) Find the largest integer $n > 3$ such that any integer (strictly) between 1 and n and relatively prime to n is a prime number.

b) Determine the largest odd integer $n > 3$ such that any odd integer (strictly) between 1 and n and relatively prime to n is a prime number.

Proof. a) Let n be such an integer and let p_1, p_2, \dots be the increasing sequence of primes. Let m be the largest positive integer such that $p_m^2 < n$, so $n \leq p_{m+1}^2$. Assume that $m \geq 4$, then $p_1 \dots p_m > p_{m+1}^2 \geq n$ by Bonse's inequality, hence n is relatively prime to one of the primes p_1, \dots, p_m , say with p_j . Then $p_j^2 \leq p_m^2 < n$ is relatively prime to n and not a prime, contradiction. Thus $m \leq 3$ and so $n \leq p_4^2 = 49$. Assume that $n > 25$. Then n cannot be relatively prime to 4, 9, 25, hence n must be a multiple of 2, 3 and 5 and so a multiple of 30. Since $n \leq 49$, this yields $n = 30$. Conversely, the smallest composite number

relatively prime to 30 is 49, so 30 has the desired property and so it is indeed the solution of the problem.

b) The argument is similar: if n is such an integer and $p_m^2 < n \leq p_{m+1}^2$, then we cannot have $m \geq 5$: otherwise $p_2 \dots p_m > p_{m+1}^2$ by Bonse's inequality and as above n is relatively prime to some p_j^2 with $2 \leq j \leq m$, contradicting the hypothesis. Thus $m \leq 4$ and $n \leq p_5^2 = 121$. Assuming that $n > 49$, we see that n cannot be relatively prime to 9, 25, 49 and so n is a multiple of $3 \cdot 5 \cdot 7 = 105$. Since $n \leq 121$, this yields $n = 105$. Conversely, the smallest odd composite number relatively prime to 105 is 121, so 105 is the solution of the problem. \square

Example 4.84. (Kolmogorov Cup) Find all odd primes p such that $1 + k(p-1)$ is prime for all $k \in \{1, 2, \dots, \frac{p-1}{2}\}$.

Proof. One checks that $p = 3$ is a solution of the problem, so assume that $p \geq 5$. Suppose that $q \leq \frac{p-1}{2}$ is a prime and that q does not divide $p-1$. Then we can find $k \in \{1, \dots, q\} \subset \{1, \dots, \frac{p-1}{2}\}$ such that $q \mid 1 + k(p-1)$, since the numbers $p-1, 2(p-1), \dots, q(p-1)$ give pairwise distinct remainders when divided by q . But then $1 + k(p-1)$ is not prime, since it is divisible by q and greater than q .

The previous paragraph shows that $p-1$ must be a multiple of all primes not exceeding $\frac{p-1}{2}$. Let $p_1 = 2, p_2 = 3, \dots$ be the sequence of primes and let m be the largest positive integer for which $p_m \leq \frac{p-1}{2}$. Then $p_1 \dots p_m \mid p-1$ by the above discussion, hence

$$p-1 \geq p_1 p_2 \dots p_m.$$

If $m \geq 4$ Bonse's inequality yields

$$p-1 > p_{m+1}^2 > \left(\frac{p-1}{2}\right)^2,$$

which contradicts the assumption that $p \geq 5$. Thus $m \leq 3$ and since $\frac{p-1}{2} < p_{m+1}$ we obtain $p < 15$. A tedious check shows that $p = 3$ and $p = 7$ are the solutions of the problem. \square

4.4 Arithmetic functions

4.4.1 Classical arithmetic functions

We will discuss in this section a few properties of some classical arithmetic functions, such as the number of divisors of a given integer, its sum of divisors, the number of prime factors of that integer, Euler's totient function, the Möbius function, etc. Before saying anything more about specific arithmetic functions, let us make clear what we mean by that:

Definition 4.85. An arithmetic function is a map $f : \mathbf{N} \rightarrow \mathbf{C}$ defined on the set of positive integers, with complex values.

Readers not comfortable with complex numbers can very well assume that all arithmetic functions take real values (as will be the case in practice). Actually, most of the time we will deal with integer-valued arithmetic functions, but it is useful to include more general functions as well (for instance, since we will often consider the quotient of two integer-valued arithmetic functions, or the square root of an arithmetic function).

From time to time it is more convenient to think of an arithmetic function f as being defined on $[1, \infty)$, by defining $f(x) = f(\lfloor x \rfloor)$ for $x \geq 1$. We will always take this convention when writing $f(x)$ for some $x \geq 1$ (not necessarily an integer) and some arithmetic function f . Note that we could have also included 0 in the domain of f , or allowed negative integers, etc.

Let us give a few classical examples of arithmetic functions, which will also allow us to introduce notation that will be used from now on constantly when dealing with arithmetic functions.

1. One of the most important arithmetic functions is Euler's totient function φ , defined by letting $\varphi(n)$ be the number of integers between 1 and n (inclusive) that are relatively prime to n . This fundamental function will be studied in more detail later on in this section. For example, $\varphi(12) = 4$, since the numbers relatively prime to 12 between 1 and 12 are 1, 5, 7, 11.
2. We define the arithmetic function τ by letting $\tau(n)$ be the number of

positive divisors of n . For instance $\tau(12) = 6$ since the divisors of 12 are 1, 2, 3, 4, 6, 12.

3. The function σ is defined by letting $\sigma(n)$ be the sum of the positive divisors of n . For example $\sigma(12) = 28$.
4. The functions ω and Ω are defined by: $\omega(n)$ is the number of different prime factors of n (with the convention that $\omega(1) = 0$); $\Omega(n)$ is the number of prime factors of n , counting multiplicities, and deciding that $\Omega(1) = 0$. In other words, if $n = p_1^{k_1} \dots p_s^{k_s}$ is the prime factorization of n , then

$$\omega(n) = s, \quad \Omega(n) = k_1 + \dots + k_s.$$

For instance $\omega(12) = 2$ and $\Omega(12) = 3$, since the prime factors of 12 are 2 (with multiplicity 2) and 3 (with multiplicity 1). Note the very useful identity

$$\Omega(ab) = \Omega(a) + \Omega(b)$$

which holds for any integers $a, b \geq 1$. On the other hand, the equality $\omega(ab) = \omega(a) + \omega(b)$ does not hold in general, but it does hold when a and b are relatively prime.

5. One of the most important arithmetic functions is π , that counts primes not exceeding n , in other words

$$\pi(n) = \sum_{p \leq n} 1$$

is the number of primes between 2 and n .

6. A very important function (studied in more detail in a later section) is the Möbius function μ . This has a rather exotic definition: $\mu(1) = 1$, $\mu(n) = 0$ if n is not squarefree (i.e. if there is a prime p such that $p^2 \mid n$) and $\mu(p_1 p_2 \dots p_k) = (-1)^k$ for distinct prime numbers p_1, \dots, p_k . In other words

$$\mu(n) = (-1)^{\omega(n)} \quad \text{if } \omega(n) = \Omega(n), \quad \mu(n) = 0 \quad \text{otherwise.}$$

7. For any prime p one can define an arithmetic function v_p by letting $v_p(n)$ be the exponent of p in the prime factorization of n . These functions play a key role in the study of primes and congruences, and chapter 5 will be devoted to them.
8. For each $k \geq 2$, define a function r_k by setting $r_k(n)$ to be the number of k -tuples of integers (x_1, \dots, x_k) such that $n = x_1^2 + \dots + x_k^2$. These functions also play a very important role in arithmetic, and we will find later on an explicit formula for $r_2(n)$. Finding $r_3(n)$ is a much more difficult problem.
9. If f is an arithmetic function, one can create two new arithmetic functions by setting

$$F(n) = \sum_{k=1}^n f(k), \quad G(n) = \sum_{d|n} f(d).$$

Many difficult problems and theorems in analytic number theory are concerned with the behavior of the functions F and G when f is one of the functions introduced above.

10. More generally, if f and g are arithmetic functions, we can define a new arithmetic function $f * g$ (called the convolution product of f and g) by

$$f * g(n) = \sum_{d|n} f(d)g\left(\frac{n}{d}\right),$$

the sum being taken over the positive divisors d of n . For instance $\tau = 1 * 1$, where 1 is the arithmetic function sending every n to 1 , and $\sigma = 1 * \text{id}$ where id is the identity function, sending every n to n . We leave it to the reader to check that $f * g = g * f$ and $(f * g) * h = f * (g * h)$ for any arithmetic functions f, g, h .

Before moving on to more theoretical results, let us discuss a few problems that involve some of the previously introduced functions. The simple observation that when d runs over the positive divisors of n , so does $\frac{n}{d}$ is a source of many identities in number theory. We invoke this very simple but rather useful observation to give a few more practical examples.

Example 4.86. Prove that for all $n > 1$ we have

$$\prod_{d|n} d = n^{\frac{\tau(n)}{2}}.$$

Proof. If $1 = d_1 < d_2 < \dots < d_k = n$ are the positive divisors of n , then so are $\frac{n}{d_k} < \frac{n}{d_{k-1}} < \dots < \frac{n}{d_1}$. Hence $d_1 d_k = d_2 d_{k-1} = \dots = d_k d_1 = n$, and multiplying these equalities yields

$$(d_1 d_2 \dots d_k)^2 = n^k = n^{\tau(n)}.$$

The result follows, since $\prod_{d|n} d = d_1 d_2 \dots d_k$. \square

Example 4.87. Show that if $n + 1$ is a multiple of 24, then $\sigma(n)$ is a multiple of 24.

Proof. First, we observe that n is not a square, since otherwise $n + 1$ would not even be a multiple of 3. Thus the positive divisors of n can be partitioned into pairs (a, b) , with $ab = n$. Since $\sigma(n)$ is the sum of the elements of these pairs, it is enough to prove that $a + b \equiv 0 \pmod{24}$ whenever $ab = n$. Now, $ab \equiv -1 \pmod{24}$, hence a and b are odd and relatively prime to 3. But if x is an odd integer relatively prime to 3, we have $x^2 \equiv 1 \pmod{24}$. Indeed, $x^2 \equiv 1 \pmod{3}$ is immediate and $x^2 \equiv 1 \pmod{8}$ is classical. Thus $ab \equiv -1 \pmod{24}$ implies $a \equiv ab^2 \equiv -b \pmod{24}$, which is the desired result $a + b \equiv 0 \pmod{24}$. \square

Example 4.88. (IMO 2002) Let $n \geq 2$ be a positive integer with divisors $1 = d_1 < d_2 < \dots < d_k = n$. Prove that $d_1 d_2 + d_2 d_3 + \dots + d_{k-1} d_k$ is less than n^2 , and determine when it is a divisor of n^2 .

Proof. Since $d_i \cdot d_{k+1-i} = n$, we can write

$$d_1 d_2 + d_2 d_3 + \dots + d_{k-1} d_k = \frac{n}{d_k} \frac{n}{d_{k-1}} + \frac{n}{d_{k-1}} \frac{n}{d_{k-2}} + \dots + \frac{n}{d_2} \frac{n}{d_1}.$$

It suffices therefore to prove that

$$\frac{1}{d_1 d_2} + \frac{1}{d_2 d_3} + \dots + \frac{1}{d_{k-1} d_k} < 1.$$

However, we have $d_i \geq i$, since the sequence d_1, \dots, d_k is strictly increasing. Hence

$$\begin{aligned} \frac{1}{d_1 d_2} + \frac{1}{d_2 d_3} + \dots + \frac{1}{d_{k-1} d_k} &\leq \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \dots + \frac{1}{(k-1)k} \\ &= 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \dots + \frac{1}{k-1} - \frac{1}{k} = 1 - \frac{1}{k} < 1. \end{aligned}$$

Now, suppose that $S = d_1 d_2 + \dots + d_{k-1} d_k$ divides n^2 and observe that

$$1 < \frac{n^2}{S} \leq \frac{n^2}{d_{k-1} d_k} = \frac{n}{d_{k-1}} = d_2.$$

Now by definition d_2 is the smallest prime divisor of n , which is also the smallest prime divisor of n^2 . On the other hand, the above inequality shows that $\frac{n^2}{S}$ is a proper divisor of n^2 which does not exceed d_2 . It follows that $\frac{n^2}{S} = d_2$ and $S = d_{k-1} d_k$, that is $k = 2$. Hence $n = d_2$ is a prime. Conversely, if n is a prime, then $S = n$ divides n^2 . Thus S divides n^2 if and only if n is a prime number. \square

The next problems are related to the function Ω .

Example 4.89. (China TST 2013) For a positive integer $N > 1$ with prime factorization $N = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$, we define $\Omega(N) = \alpha_1 + \alpha_2 + \dots + \alpha_k$. Let a_1, a_2, \dots, a_n be positive integers and let $f(x) = (x + a_1)(x + a_2) \dots (x + a_n)$. Prove that if $\Omega(f(k))$ is even for all positive integers k , then n is even.

Proof. Since $\Omega(ab) = \Omega(a) + \Omega(b)$ for all integers $a, b > 1$, it follows from the hypothesis of the problem that $\Omega(f(x_1) \dots f(x_k))$ is even for all positive integers x_1, \dots, x_k . One easily checks that

$$f(1) \cdot \prod_{i=1}^n f(a_i + 2) = 2^n \cdot \prod_{i=1}^n (a_i + 1)^2 \prod_{1 \leq i < j \leq n} (a_i + a_j + 2)^2$$

We deduce that $\Omega(2^n) = n$ is even. \square

Example 4.90. (Romanian Masters in Mathematics 2011) Given a positive integer n with prime factorization $n = \prod_{i=1}^s p_i^{\alpha_i}$, let $\lambda(n) = (-1)^{\alpha_1 + \dots + \alpha_s}$. Prove that:

a) There are infinitely many positive integers n such that

$$\lambda(n) = \lambda(n+1) = 1.$$

b) For infinitely many n we have $\lambda(n) = \lambda(n+1) = -1$.

Proof. We start by observing that $\lambda(mn) = \lambda(m) \cdot \lambda(n)$ for all positive integers m, n , and that $\lambda(n^2) = 1$ for any positive integer n .

a) Note that $\lambda(9) = \lambda(10) = 1$, so there are certainly positive integers n such that $\lambda(n) = \lambda(n+1) = 1$. Assume that there are only finitely many such n , so there is $N > 1$ such that if $n > N$ then one of the numbers $\lambda(n)$ and $\lambda(n+1)$ is different from 1. If $a > N+1$, then $\lambda(a^2 - 1)$ cannot be 1, since $\lambda(a^2) = 1$. Thus $\lambda(a^2 - 1) = -1$ and so $\lambda(a-1) + \lambda(a+1) = 0$ for $a > N+1$. In particular $\lambda(a) = -\lambda(a+2) = \lambda(a+4)$ for $a > N+1$. If $x > N+1$, we deduce that

$$1 = \lambda(4x^2) = \lambda(4x^2 + 4) = \dots = \lambda(4x^2 + 4x) = \lambda((2x+1)^2 - 1) = -1,$$

a contradiction.

b) It is again not difficult to find explicitly one such n , since $\lambda(2) = \lambda(3) = -1$. Assume that there are only finitely many such n , thus there is $N > 1$ such that for $n > N$ at least one of the numbers $\lambda(n)$ and $\lambda(n+1)$ is not -1 . Take $k > N+1$ such that $\lambda(2k+1) = -1$, for instance $k = \frac{p-1}{2}$ with $p > 2N+3$ a prime. Then $\lambda(2k) = 1$ and so $\lambda(k) = -1$. But then $\lambda(k+1) = 1$ and so $\lambda(2k+2) = -1 = \lambda(2k+1)$, a contradiction. \square

Remark 4.91. The problem can be also easily solved using the Pell equation. The equation $x^2 - 6y^2 = 1$ solves part a): if (x, y) is a solution of the equation, then clearly

$$1 = \lambda(x^2) = \lambda(6y^2) = \lambda(x^2 - 1).$$

For the second part, we use the equation $3x^2 - 2y^2 = 1$, which also has infinitely many solutions.

Example 4.92. (IMO Shortlist 2009) A positive integer N is called balanced if $N = 1$ or if N can be written as a product of an even number of not necessarily distinct primes. Let a, b be positive integers and let $P(x) = (x + a)(x + b)$ for each positive integer x .

(a) Prove that there exist distinct positive integers a and b such that all numbers $P(1), P(2), \dots, P(50)$ are balanced.

(b) Prove that if $P(n)$ is balanced for all positive integers n , then $a = b$.

Proof. Let $\Omega(n)$ be the number of prime divisors of n , counted with multiplicities. Then n is balanced if and only if $\Omega(n)$ is even. We have already seen that $\Omega(ab) = \Omega(a) + \Omega(b)$ for all positive integers a, b . Thus, $\Omega(a)$ and $\Omega(b)$ have the same parity if and only if ab is balanced.

a) Our aim is to prove the existence of a, b such that $\Omega(a + i)$ and $\Omega(b + i)$ have the same parity for all $1 \leq i \leq 50$. This is a simple application of the pigeonhole principle: for each positive integer a consider the sequence $(x_1(a), \dots, x_{50}(a))$, where $x_i(a)$ is the remainder of $\Omega(a + i)$ when divided by 2. Since there are infinitely many positive integers and only finitely many sequences of length 50 with entries in $\{0, 1\}$, two positive integers a, b will have the same associated sequence. This is just another way of saying that $\Omega(a + i)$ and $\Omega(b + i)$ have the same parity for all $1 \leq i \leq 50$, so we are done.

b) Suppose that $a \neq b$ and, without loss of generality, that $a < b$. By assumption $\Omega(n + a) \equiv \Omega(n + b) \pmod{2}$ for all $n \geq 1$, thus $\Omega(k) \equiv \Omega(k + b - a) \pmod{2}$ for all $k \geq 1$. It follows that $\Omega(k) \equiv \Omega(k + j(b - a)) \pmod{2}$ for all $k, j \geq 1$. In particular

$$\Omega(b(b - a)) \equiv \Omega(b(b - a) + b(b - a)) = \Omega(2b(b - a)) = 1 + \Omega(b(b - a)) \pmod{2},$$

which is certainly absurd. \square

4.4.2 Multiplicative functions

A very important class of arithmetic functions is that of multiplicative (respectively totally multiplicative) functions, which we define as follows:

Definition 4.93. An arithmetic function f is called multiplicative (respectively totally multiplicative) if $f(mn) = f(m)f(n)$ for all relatively prime positive integers m, n (respectively for all positive integers m, n).

Let us make a few simple remarks about multiplicative functions. First, note that any totally multiplicative function is multiplicative, but the converse is false. Also, note that if f is a multiplicative function, then

$$f(n) = f(n \cdot 1) = f(n)f(1)$$

for all positive integers n , thus either f vanishes identically or $f(1) = 1$. Thus all interesting multiplicative functions f satisfy $f(1) = 1$. Secondly, if f is a multiplicative function, then f is uniquely determined by its values on prime powers, since any positive integer can be written as a product of powers of primes, and

$$f(p_1^{k_1} \dots p_n^{k_n}) = f(p_1^{k_1}) \dots f(p_n^{k_n})$$

for all pairwise distinct primes p_1, \dots, p_n and all nonnegative integers k_1, \dots, k_n . A very useful consequence of this observation is that if we are asked to prove that two multiplicative functions f, g are equal, then it suffices to check that they agree on prime powers (which is usually much easier to check in practice!).

Many important arithmetic functions are multiplicative. The next simple theorem establishes the multiplicative character of the functions τ and σ , by giving explicit formulae for $\tau(n)$ and $\sigma(n)$ in terms of the prime factorization of n . These formulae are very important when dealing with these functions.

Theorem 4.94. *If $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_m^{\alpha_m}$ is the prime factorization of $n > 1$, then*

$$\tau(n) = (\alpha_1 + 1)(\alpha_2 + 1) \dots (\alpha_m + 1)$$

and

$$\sigma(n) = \prod_{i=1}^m (1 + p_i + \dots + p_i^{\alpha_i}) = \frac{p_1^{\alpha_1+1} - 1}{p_1 - 1} \cdot \dots \cdot \frac{p_m^{\alpha_m+1} - 1}{p_m - 1}.$$

Proof. The fundamental theorem of arithmetic allows us to describe all positive divisors of $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_m^{\alpha_m}$. Namely, they are exactly all numbers $p_1^{\beta_1} \dots p_m^{\beta_m}$ for some $\beta_1 \in \{0, 1, \dots, \alpha_1\}, \dots, \beta_m \in \{0, 1, \dots, \alpha_m\}$ (and two such divisors are equal if and only if the corresponding m -tuples $(\beta_1, \dots, \beta_m)$ and $(\beta'_1, \dots, \beta'_m)$ are equal). Since β_i can take $\alpha_i + 1$ possible values, the formula for $\tau(n)$ is clear. For $\sigma(n)$ we obtain

$$\sigma(n) = \sum_{0 \leq \beta_1 \leq \alpha_1} \sum_{0 \leq \beta_2 \leq \alpha_2} \dots \sum_{0 \leq \beta_m \leq \alpha_m} p_1^{\beta_1} \dots p_m^{\beta_m}$$

$$= \left(\sum_{0 \leq \beta_1 \leq \alpha_1} p_1^{\beta_1} \right) \cdot \dots \cdot \left(\sum_{0 \leq \beta_m \leq \alpha_m} p_m^{\beta_m} \right)$$

and the result follows using the identity

$$1 + x + \dots + x^n = \frac{x^{n+1} - 1}{x - 1}. \quad \square$$

The next problems illustrate the use of the previous explicit formulae for the τ function.

Example 4.95. Prove that $\tau(n)$ is odd if and only if n is a perfect square.

Proof. If $n = p_1^{a_1} \dots p_k^{a_k}$ is the prime factorization of n , then

$$\tau(n) = (a_1 + 1)(a_2 + 1) \dots (a_k + 1)$$

is odd if and only if each factor $a_i + 1$ is odd, that is if and only if each a_i is even. This is of course equivalent to n being a square. \square

Example 4.96. (Belarus 1999) Let a, b be positive integers such that the product of all positive divisors of a equals the product of all positive divisors of b . Prove that $a = b$.

Proof. By assumption and example 4.86 we have $a^{\tau(a)} = b^{\tau(b)}$. This immediately implies that a and b have the same prime factors, call them p_1, \dots, p_k . Let $a = p_1^{x_1} \dots p_k^{x_k}$ and $b = p_1^{y_1} \dots p_k^{y_k}$ for some positive integers $x_1, \dots, x_k, y_1, \dots, y_k$. The equality $a^{\tau(a)} = b^{\tau(b)}$ forces $x_i \tau(a) = y_i \tau(b)$ for all i . Let

$$u = \frac{\tau(a)}{\gcd(\tau(a), \tau(b))} \text{ and } v = \frac{\tau(b)}{\gcd(\tau(a), \tau(b))},$$

so that $\gcd(u, v) = 1$ and $ux_i = vy_i$. We deduce that $y_i = uz_i$ and $x_i = vz_i$ for some positive integers z_i . Thus

$$a = (p_1^{z_1} \dots p_k^{z_k})^u, \quad b = (p_1^{z_1} \dots p_k^{z_k})^v.$$

Clearly, if $u > v$ then

$$\tau(a) = (1 + uz_1) \dots (1 + uz_k) > (1 + vz_1) \dots (1 + vz_k) = \tau(b)$$

and so $a^{\tau(a)} > b^{\tau(b)}$. Similarly we cannot have $u < v$, thus $u = v$, $x_i = y_i$ for all i and finally $a = b$. \square

Example 4.97. Prove that for all $n > 1$ we have

$$\tau((n-1)!) \geq \frac{\tau(n!)}{2}.$$

Proof. If n is a prime, then n and $(n-1)!$ are relatively prime and so the proposed inequality is an equality. Assume from now on that n is composite and write

$$n = p_1^{a_1} \dots p_k^{a_k}$$

for its prime factorization (note that $p_i < n$ for all i). Write

$$(n-1)! = p_1^{b_1} \dots p_k^{b_k} \cdot q_1^{c_1} \dots q_s^{c_s},$$

where q_i are the primes not exceeding $n-1$ and not belonging to $\{p_1, \dots, p_k\}$. Then

$$\frac{\tau(n!)}{\tau((n-1)!)} = \prod_{i=1}^k \frac{a_i + b_i + 1}{b_i + 1} = \prod_{i=1}^k \left(1 + \frac{a_i}{b_i + 1}\right).$$

We need to prove that this expression is ≤ 2 . Note that since $p_i \mid n$, the numbers $p_i, 2p_i, \dots, \left(\frac{n}{p_i} - 1\right)p_i$ appear in the product defining $(n-1)!$, thus

$$\frac{a_i}{b_i + 1} \leq \frac{a_i}{\frac{n}{p_i}} = \frac{a_i p_i}{n}.$$

Letting $x_i = p_i^{a_i}$, we have $x_i \geq 2$ for all i , $n = x_1 \dots x_k$ and x_1, \dots, x_k are pairwise distinct integers. Moreover, we clearly have

$$x_i = p_i \cdot p_i^{a_i-1} \geq p_i \cdot 2^{a_i-1} \geq a_i p_i.$$

It is thus sufficient to prove that

$$\prod_{i=1}^k \left(1 + \frac{x_i}{n}\right) \leq 2.$$

This is clear if $k = 1$. For $k \geq 2$ it follows by an easily noting that the inequality

$$\left(1 + \frac{x}{n}\right) \left(1 + \frac{y}{n}\right) \leq 1 + \frac{xy}{n}$$

rearranges to $\frac{1}{x} + \frac{1}{y} + \frac{1}{n} \leq 1$. Since for $k \geq 2$, we have $x_1 \geq 2$, $x_2 \geq 3$, and $n \geq 6$, this inequality holds for x and y any nonempty product of x_i 's. Iterating this gives

$$\prod_{i=1}^k \left(1 + \frac{x_i}{n}\right) \leq 1 + \frac{x_1 x_2 \cdots x_k}{n} = 2. \quad \square$$

Example 4.98. (China TST 2015) For $n > 1$ define

$$f(n) = \tau(n!) - \tau((n-1)!).$$

Prove that there are infinitely many composite numbers n such that for all $1 < m < n$ we have $f(m) < f(n)$.

Proof. We try some of the simplest possible composite numbers, namely $n = 2p$ with $p > 2$ a prime. We will prove that they are all solutions of the problem. Let us compute first $f(2p) = \tau((2p)!) - \tau((2p-1)!)$. Note that $(2p-1)!$ is divisible by p exactly once, so we can write $(2p-1)! = px$ with x relatively prime to p . Then $(2p)! = 2p^2x$ and so

$$f(2p) = \tau(2p^2x) - \tau(px) = \tau(p^2)\tau(2x) - \tau(p)\tau(x) > 3\tau(x) - 2\tau(x) = \tau(x),$$

the inequality being a consequence of the fact that $\tau(2x) > \tau(x)$. It is thus enough to prove that for each $m \in \{2, 3, \dots, 2p-1\}$ we have $f(m) \leq \tau(x)$. By example 4.97 we know that

$$f(m) \leq \frac{\tau(m!)}{2} \leq \frac{\tau((2p-1)!)}{2} = \frac{\tau(px)}{2} = \tau(x),$$

thus we are done. \square

We will give now another argument for the multiplicative character of the functions τ and σ , since this argument applies in many other situations. Note that

$$\tau(n) = \sum_{d|n} 1, \quad \sigma(n) = \sum_{d|n} d$$

and that the constant function 1 and the identity function are obviously multiplicative. The next theorem immediately implies that τ and σ are multiplicative. Before stating this theorem, we recall that if f, g are arithmetic functions, the convolution product $f * g$ of f and g is defined by

$$(f * g)(n) = \sum_{d|n} f(d)g\left(\frac{n}{d}\right).$$

Theorem 4.99. *The convolution product of two multiplicative functions is a multiplicative function. In particular, if f is multiplicative, then the function F defined by*

$$F(n) = \sum_{d|n} f(d)$$

is also multiplicative.

Proof. Suppose that f and g are multiplicative and let m, n be relatively prime positive integers. Then each positive divisor d of mn can be uniquely written $d = d_1 d_2$, with d_1, d_2 positive divisors of m and n respectively. This follows easily from the fundamental theorem of arithmetic and from Gauss' lemma. Hence we can write

$$f * g(mn) = \sum_{d|mn} f(d)g\left(\frac{mn}{d}\right) = \sum_{d_1|m, d_2|n} f(d_1 d_2)g\left(\frac{m}{d_1} \cdot \frac{n}{d_2}\right).$$

Now, note that since $\gcd(m, n) = 1$, we also have

$$\gcd(d_1, d_2) = 1 \text{ and } \gcd\left(\frac{m}{d_1}, \frac{n}{d_2}\right) = 1.$$

Thus using the fact that f and g are multiplicative we obtain

$$f * g(mn) = \sum_{d_1|m, d_2|n} f(d_1)f(d_2)g\left(\frac{m}{d_1}\right)g\left(\frac{n}{d_2}\right)$$

$$= \sum_{d_1|m} f(d_1)g\left(\frac{m}{d_1}\right) \cdot \sum_{d_2|n} f(d_2)g\left(\frac{n}{d_2}\right) = f * g(m) \cdot f * g(n),$$

proving that $f * g$ is multiplicative. \square

Example 4.100. (Liouville's theorem) Prove that for all positive integers n ,

$$\left(\sum_{d|n} \tau(d)\right)^2 = \sum_{d|n} \tau(d)^3.$$

Proof. Both sides are multiplicative functions of n by the previous theorem, hence it suffices to prove the equality when n is a power of a prime p , say $n = p^k$. Then

$$\sum_{d|n} \tau(d) = \sum_{j=0}^k \tau(p^j) = \sum_{j=1}^{k+1} j = \frac{(k+1)(k+2)}{2}$$

and

$$\sum_{d|n} \tau(d)^3 = \sum_{j=1}^{k+1} j^3 = \frac{(k+1)^2(k+2)^2}{4}.$$

The result follows. \square

We end this section with some miscellaneous problems in which the concept of multiplicative function plays a crucial role.

Example 4.101. (Balkan Mathematical Olympiad 1991) Prove that there is no bijection $f : \mathbf{N} \rightarrow \{0, 1, 2, \dots\}$ such that for all $m, n \in \mathbf{N}$

$$f(mn) = f(m) + f(n) + 3f(m)f(n).$$

Proof. Assuming that such a bijection f exists, define $g(n) = 3f(n) + 1$ and let S be the set of positive integers congruent to 1 mod 3. Then $g : \mathbf{N} \rightarrow S$ is a bijection such that $g(mn) = g(m)g(n)$ for any $m, n \in \mathbf{N}$, i.e. g is totally multiplicative, in particular $g(1) = 1$. Let $p, q, r \in \mathbf{N}$ be such that $g(p) = 4$, $g(q) = 10$ and $g(r) = 25$. Since any of the numbers 4, 10 and 25 is not a product

of two numbers from the set $S \setminus \{1\}$ and since g is totally multiplicative, it follows that p , q and r are distinct prime numbers. On the other hand,

$$g(pr) = g(p)g(r) = 10^2 = g^2(q) = g(q^2)$$

and so $pr = q^2$, a contradiction. \square

Example 4.102. (Turkey 1995) Find all surjective functions $f : \mathbf{N} \rightarrow \mathbf{N}$ such that for all $m, n \in \mathbf{N}$ we have $m \mid n$ if and only if $f(m) \mid f(n)$.

Proof. Note that f is injective, since $f(m) = f(n)$ forces $m \mid n$ and $n \mid m$, thus $m = n$. Next, $f(1) \mid f(n)$ for all $n \geq 1$ and since there is n such that $f(n) = 1$, we deduce that $f(1) = 1$.

Let m, n be relatively prime positive integers. Then $f(m)$ and $f(n)$ are relatively prime: if they had a common divisor $d > 1$, then $d = f(k)$ for some $k > 1$ and then k divides both m and n , a contradiction. Next, since $f(m)$ and $f(n)$ both divide $f(mn)$, we deduce that $f(m)f(n) \mid f(mn)$. On the other hand, $f(m)f(n) = f(c)$ for some $c \geq 1$, and c is a multiple of m and n , thus a multiple of mn . But $f(c) \mid f(mn)$, thus $c \mid mn$ and finally $c = mn$. In other words, $f(mn) = f(m)f(n)$ when m, n are relatively prime and so if $n = p_1^{k_1} \dots p_r^{k_r}$ is the prime factorization of n , then

$$f(n) = f(p_1^{k_1}) \dots f(p_r^{k_r}).$$

It remains thus to understand $f(p^k)$ when p is a prime and $k \geq 1$. Note that $f(p) > 1 = f(1)$, since f is injective, and $f(p)$ has no proper divisor: if d was such a divisor, then $d = f(c)$ and c would be a proper divisor of p , impossible. Thus $f(p)$ is also a prime. Conversely, if $f(n)$ is a prime for some n , then n is a prime (same argument as above). Thus the restriction of f to the set of prime numbers is a permutation of this set.

Finally, we will prove that $f(p^k) = f(p)^k$ for any prime p and any $k \geq 1$, by induction on k . Assume that $f(p^j) = f(p)^j$ for $1 \leq j \leq k-1$. Then $f(p^k)$ is divisible by $f(p)^{k-1}$ and its divisors are precisely $f(c)$ with $c \mid p^k$, that is the numbers $1, f(p), \dots, f(p)^{k-1}, f(p^k)$. We deduce from this that $f(p^k) = f(p)^{k-1} \cdot f(p) = f(p)^k$ and the inductive step is finished.

The previous discussion shows that there is a permutation $(a_p)_p$ of the set of prime numbers such that

$$f(n) = \prod_{p|n} a_p^{v_p(n)}.$$

Conversely, it is clear that any such function is a solution of the problem. \square

Example 4.103. (IMO Shortlist 1996) Find a bijection $f: \{0, 1, 2, \dots\} \rightarrow \{0, 1, 2, \dots\}$ that satisfies

$$f(3mn + m + n) = 4f(m)f(n) + f(m) + f(n)$$

for all $m, n \geq 0$.

Proof. Note that the condition can be written

$$f\left(\frac{(3m+1)(3n+1)-1}{3}\right) = \frac{(4f(m)+1)(4f(n)+1)-1}{4}.$$

Letting $A = \{3k+1 \mid k \geq 0\}$, the previous relation suggests defining a function $h: A \rightarrow \{1, 2, \dots\}$ by

$$h(x) = 4f\left(\frac{x-1}{3}\right) + 1.$$

The problem is then equivalent to constructing a bijection h between A and the set $B = \{4k+1 \mid k \geq 0\}$ such that $h(mn) = h(m)h(n)$ for all $m, n \in A$.

We set $h(1) = 1$ and consider the set U of all primes of form $3k-1$, the set V of all primes of form $3k+1$, the set X of all primes of form $4k-1$ and finally the set Y of all primes of form $4k+1$. By Dirichlet's theorem each of the sets U, V, X, Y is infinite. (An elementary proof of this for U and X was given in example 4.56. For V and Y an elementary proof will be given in example 5.31.) Thus we can construct a bijection h_1 between U and X and a bijection h_2 between V and Y (to do so, enumerate in increasing order the elements $u_1 < u_2 < \dots$ and $x_1 < x_2 < \dots$ of U , respectively X , and map u_1 to x_1 , u_2 to x_2, \dots). If $n \in A$ and

$$n = \prod_{i=1}^k u_i^{a_i} \cdot \prod_{i=1}^l v_i^{b_i}$$

is the prime factorization of n , define

$$h(n) = \prod_{i=1}^k h_1(u_i)^{a_i} \cdot \prod_{i=1}^l h_2(v_i)^{b_i}.$$

Note that $h(n) \in B$, since $\sum_{i=1}^k a_i$ is even (because $n \equiv 1 \pmod{3}$ and $u_i \equiv -1 \pmod{3}$, while $v_i \equiv 1 \pmod{3}$) and $h_1(u_i) \equiv -1 \pmod{4}$, while $h_2(v_i) \equiv 1 \pmod{4}$ for all i . One can construct an inverse h^{-1} of h using the inverses of h_1 and h_2 on X and Y , using exactly the same recipe and arguments as above. \square

Example 4.104. (IMO 1998) Consider all functions $f : \mathbf{N} \rightarrow \mathbf{N}$ such that

$$f(n^2 f(m)) = m f(n)^2$$

for all $m, n \in \mathbf{N}$. Find the least possible value of $f(1998)$.

Proof. Let f be such a function and define $a = f(1)$. Since $f(f(m)) = a^2 m$ and $f(an^2) = f(n)^2$ for all m, n (set $n = 1$ and $m = 1$ in the given relation), we obtain

$$f(m)^2 f(n)^2 = f(m)^2 f(an^2) = f(m^2 f(f(an^2))) = f(m^2 a^3 n^2) = f(amn)^2,$$

i.e. $f(m)f(n) = f(amn)$. In particular, $f(am) = af(m)$ and therefore

$$af(mn) = f(m)f(n).$$

An immediate induction then shows that $f(n)^k = a^{k-1} f(n^k)$ for all k , thus $a^{k-1} \mid f(n)^k$ for all k . If p is a prime factor of a and if α, β are the exponents of p in the prime factorization of a , respectively $f(n)$, we obtain $(k-1)\alpha \leq k\beta$ for all $k \geq 1$, thus $\alpha \leq \beta$. It follows that a divides $f(n)$ for all $n \in \mathbf{N}$, hence the function

$$g : \mathbf{N} \rightarrow \mathbf{N}, \quad g(n) = \frac{f(n)}{a}$$

is well-defined and satisfies

$$g(mn) = g(m)g(n) \quad \text{and} \quad g(g(m)) = m$$

for all $m, n \in \mathbf{N}$. In particular, g is bijective, and moreover g maps prime numbers to prime numbers. Indeed, if p is a prime and $g(p) = ab$ for some integers $a, b > 1$, then $p = g(g(p)) = g(a)g(b)$, thus $g(a) = 1 = g(1)$ or $g(b) = 1 = g(1)$, contradicting the injectivity of g . Letting P be the set of prime numbers, we obtain that $g : P \rightarrow P$ is an involution, i.e. $g(g(p)) = p$. Conversely, given an involution g of P and $a \in \mathbf{N}$, one obtains a map f as in the statement of the problem by defining $f(n) = ag(n)$, where $g(1) = 1$ and

$$g(n) = \prod_{i=1}^k g(p_i)^{\alpha_i}$$

if $n = \prod_{i=1}^k p_i^{\alpha_i}$ is the prime factorization of $n > 1$.

Finally, observe that since $g(2), g(3)$ and $g(37)$ are different prime numbers, we have

$$g(2)g(3)^3g(37) \geq 3 \cdot 2^3 \cdot 5 = 120,$$

hence

$$f(1998) = f(2 \cdot 3^3 \cdot 37) = f(1)g(2)g(3)^3g(37) \geq 120.$$

In order to see that this lower bound is attained, set

$$a = f(1) = 1, g(2) = 3, g(3) = 2, g(5) = 37, g(37) = 5$$

and $g(p) = p$ for all prime numbers $p \neq 2, 3, 5, 37$. Then $g(g(p)) = p$ for all $p \in P$ and as we said above these data determine uniquely a function $f : \mathbf{N} \rightarrow \mathbf{N}$ with the desired properties. Thus the answer of the problem is 120. \square

4.4.3 Euler's phi function

In this section we study in more detail the fundamental totient function $\varphi : \mathbf{N} \rightarrow \mathbf{N}$. Recall that $\varphi(n)$ is the number of integers between 1 and n (inclusive) that are relatively prime to n . The map φ is called *Euler's totient function* or *Euler's phi function*, while an integer $a \in \{1, 2, \dots, n\}$ which is relatively prime to n is called a *totative* of n .

Clearly $\varphi(1) = 1$ and $\varphi(p) = p - 1$ for any prime p , since the totatives of p are $1, 2, \dots, p - 1$. More generally, if $n \geq 1$ and p is a prime, then the totatives

of p^n are the numbers in $\{1, 2, \dots, p^n\}$ which are not divisible by p . Since there are p^{n-1} multiples of p in $\{1, 2, \dots, p^n\}$, it follows that

$$\varphi(p^n) = p^n - p^{n-1} = p^{n-1}(p - 1) = p^n \left(1 - \frac{1}{p}\right).$$

We will now explain how to find a closed formula for $\varphi(n)$, using a combinatorial argument based on the following very useful result (we denote by $|A|$ the number of elements of a finite set A).

Proposition 4.105. (*inclusion-exclusion principle*) *For any family of finite subsets A_1, \dots, A_k of a set X we have*

$$\left| \bigcup_{i=1}^k A_i \right| = \sum_{i=1}^k |A_i| - \sum_{1 \leq i < j \leq k} |A_i \cup A_j| + \dots + (-1)^{k-1} |A_1 \cap \dots \cap A_k|.$$

Proof. If $B \subset X$ is a subset and $x \in X$, let $1_{x \in B}$ be equal to 1 if $x \in B$ and 0 otherwise. Then clearly $|B| = \sum_{x \in X} 1_{x \in B}$ if $B \subset X$ is finite and $1_{x \in B_1 \cap \dots \cap B_d} = 1_{x \in B_1} \cdot \dots \cdot 1_{x \in B_d}$ for all subsets B, B_1, \dots, B_d of X . Let R be the right-hand side of the equality we want to establish. Then using the above observations we obtain

$$R = \sum_{x \in X} \left[\sum_{i=1}^k 1_{x \in A_i} - \sum_{i < j} 1_{x \in A_i} \cdot 1_{x \in A_j} + \dots + (-1)^{k-1} 1_{x \in A_1} \cdot \dots \cdot 1_{x \in A_k} \right].$$

Using the identity

$$\sum_{i=1}^k z_i - \sum_{i < j} z_i z_j + \dots + (-1)^{k-1} z_1 \dots z_k = 1 - (1 - z_1) \dots (1 - z_k)$$

we obtain

$$R = \sum_{x \in X} \left(1 - \prod_{i=1}^k (1 - 1_{x \in A_i}) \right).$$

On the other hand, it is clear that for all $x \in X$ we have

$$1 - \prod_{i=1}^k (1 - 1_{x \in A_i}) = 1_{x \in A_1 \cup \dots \cup A_k},$$

thus

$$R = \sum_{x \in X} 1_{x \in A_1 \cup \dots \cup A_k} = |A_1 \cup \dots \cup A_k|,$$

as needed. □

We are now ready to prove the following crucial theorem.

Theorem 4.106. *For all $n > 1$ we have*

$$\varphi(n) = n \cdot \prod_{p|n} \left(1 - \frac{1}{p}\right),$$

the product being taken over all prime divisors p of n , without multiplicities. Thus, if $n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$ is the prime factorization of n , then

$$\varphi(n) = p_1^{\alpha_1-1} \dots p_k^{\alpha_k-1} (p_1 - 1) \dots (p_k - 1).$$

Proof. Let $n > 1$ and let $n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$ be the prime factorization of n . Then an integer $a \in \{1, 2, \dots, n\}$ is a totative of n if and only if a is not divisible by any of the numbers p_1, \dots, p_k . Equivalently, if A_i is the set of multiples of p_i among $1, 2, \dots, n$, then the set of totatives of n is precisely the complement of $\bigcup_{i=1}^k A_i$. It follows that

$$\varphi(n) = n - \left| \bigcup_{i=1}^k A_i \right|.$$

We use the inclusion-exclusion principle to evaluate $\left| \bigcup_{i=1}^k A_i \right|$. For this, we need to evaluate the number of elements of $A_{i_1} \cap \dots \cap A_{i_r}$ for all $1 \leq r \leq k$ and all $1 \leq i_1 < \dots < i_r \leq k$. Fortunately, this is fairly easy, since $A_{i_1} \cap \dots \cap A_{i_r}$ consists of those $a \in \{1, 2, \dots, n\}$ which are multiples of p_{i_1}, \dots, p_{i_r} or equivalently multiples of $p_{i_1} p_{i_2} \dots p_{i_r}$. Thus

$$|A_{i_1} \cap \dots \cap A_{i_r}| = \frac{n}{p_{i_1} \dots p_{i_r}}.$$

We conclude that

$$\varphi(n) = n - n \cdot \sum_{i=1}^k \frac{1}{p_i} + n \cdot \sum_{1 \leq i < j \leq k} \frac{1}{p_i p_j} + \dots = n \cdot \prod_{i=1}^k \left(1 - \frac{1}{p_i}\right),$$

which finishes the proof of theorem 4.106. □

For instance, since $1000 = 2^3 5^3$, we obtain

$$\varphi(1000) = 2^2(2-1)5^2(5-1) = 400$$

and similarly $2016 = 2^5 \cdot 3^2 \cdot 7$, hence

$$\varphi(2016) = 2^4 \cdot 3 \cdot (2-1)(3-1)(7-1) = 576.$$

Example 4.107. (Komal A 240) Prove that for all $m, n \geq 1$

$$\sum_{\substack{1 \leq k \leq n \\ \gcd(k, m) = 1}} \frac{1}{k} \geq \frac{\varphi(m)}{m} \sum_{k=1}^n \frac{1}{k}.$$

Proof. Let p_1, \dots, p_s be the prime factors of m , without counting multiplicities. The inequality is equivalent to

$$\prod_{i=1}^s \frac{1}{1 - \frac{1}{p_i}} \cdot \sum_{\substack{1 \leq k \leq n \\ \gcd(k, m) = 1}} \frac{1}{k} \geq \sum_{k=1}^n \frac{1}{k},$$

or

$$\prod_{i=1}^s \left(1 + \frac{1}{p_i} + \frac{1}{p_i^2} + \dots \right) \cdot \sum_{\substack{1 \leq k \leq n \\ \gcd(k, m) = 1}} \frac{1}{k} \geq \sum_{k=1}^n \frac{1}{k}.$$

Expanding brutally the expression in the left-hand side, we obtain an infinite sum, among whose terms we have all

$$\frac{1}{p_1^{k_1} \cdot \dots \cdot p_s^{k_s} \cdot r}$$

with $k_i \geq 0$ and $1 \leq r \leq n$, $\gcd(r, m) = 1$. Since any number k between 1 and n can be written $k = p_1^{k_1} \cdot \dots \cdot p_s^{k_s} \cdot r$ with k_i and r as above, the result follows. \square

Here is another example of a nice use of the inclusion-exclusion principle.

Example 4.108. (Putnam 2015) Let q be an odd positive integer, and let N_q be the number of integers a such that $0 < a < q/4$ and $\gcd(a, q) = 1$. Prove that N_q is odd if and only if q is of the form p^k with k a positive integer and p a prime congruent to 5 or 7 modulo 8.

Proof. Let p_1, \dots, p_n be the prime divisors of q (without counting multiplicities). If A_i is the set of multiples of p_i between 0 and $\frac{q}{4}$, then

$$\begin{aligned} N_q &= \left\lfloor \frac{q}{4} \right\rfloor - |\cup_{i=1}^n A_i| = \left\lfloor \frac{q}{4} \right\rfloor - \sum_{i=1}^n |A_i| + \dots + (-1)^n |A_1 \cap \dots \cap A_n| \\ &\equiv \left\lfloor \frac{q}{4} \right\rfloor + \sum_{i=1}^n |A_i| + \dots + |A_1 \cap \dots \cap A_n| \pmod{2}. \end{aligned}$$

Note that for all i_1, \dots, i_k we have

$$|A_{i_1} \cap \dots \cap A_{i_k}| = \left\lfloor \frac{q}{4p_{i_1} \dots p_{i_k}} \right\rfloor.$$

Thus

$$N_q \equiv \left\lfloor \frac{q}{4} \right\rfloor + \sum_{i=1}^n \left\lfloor \frac{q}{4p_i} \right\rfloor + \dots + \left\lfloor \frac{q}{4p_1 \dots p_n} \right\rfloor \pmod{2}.$$

We observe next that if a, b are odd integers, then

$$\left\lfloor \frac{a}{4} \right\rfloor + \left\lfloor \frac{b}{4} \right\rfloor \equiv \left\lfloor \frac{ab}{4} \right\rfloor \pmod{2}.$$

Indeed, writing $a = 4q + r$ and $b = 4q' + r'$ with $r, r' \in \{1, 3\}$, we have

$$\left\lfloor \frac{ab}{4} \right\rfloor = 4qq' + qr' + q'r + \left\lfloor \frac{rr'}{4} \right\rfloor \equiv q + q' + \left\lfloor \frac{rr'}{4} \right\rfloor \pmod{2}$$

and it is immediate to check that $\left\lfloor \frac{rr'}{4} \right\rfloor$ is even, yielding the claim.

We conclude that

$$N_q \equiv \left\lfloor \frac{1}{4} q \cdot \prod_{i=1}^n \frac{q}{p_i} \cdot \dots \cdot \frac{q}{p_1 \dots p_n} \right\rfloor = \left\lfloor \frac{q^{2^n}}{4(p_1 \dots p_n)^{2^{n-1}}} \right\rfloor \pmod{2}.$$

If $n > 1$ then $\frac{q^{2^n}}{(p_1 \dots p_n)^{2^n - 1}}$ is the square of an odd integer and we deduce immediately that N_q is even. Assume now that $n = 1$, so that $q = p_1^k$ for some $k \geq 1$. Then N_q is odd if and only if $\left\lfloor \frac{q^2}{4p_1} \right\rfloor = \left\lfloor \frac{p_1^{2k-1}}{4} \right\rfloor$ is odd. A simple inspection shows that this happens precisely when $p \equiv 5, 7 \pmod{8}$. \square

The fundamental theorem of arithmetic combined with the formula for $\varphi(n)$ established in the previous theorem immediately yield the following result.

Corollary 4.109. *φ is a multiplicative function, that is $\varphi(mn) = \varphi(m)\varphi(n)$ for all relatively prime positive integers m, n .*

Also note another immediate consequence of the previous theorem.

Corollary 4.110. *If a, b are positive integers and $a \mid b$, then $\varphi(a) \mid \varphi(b)$.*

We end this theoretical part with an important theorem of Gauss. The proof uses the following simple but important observation.

Proposition 4.111. *For each positive divisor d of n there are precisely $\varphi(\frac{n}{d})$ integers $k \in \{1, 2, \dots, n\}$ for which $\gcd(k, n) = d$.*

Proof. We have $\gcd(n, k) = d$ if and only if $k = du$, with $u \in \{1, 2, \dots, \frac{n}{d}\}$ relatively prime to $\frac{n}{d}$. The result follows. \square

Theorem 4.112. (Gauss) *For all positive integers n we have*

$$\sum_{d \mid n} \varphi(d) = n.$$

Proof. For each $k \in \{1, 2, \dots, n\}$, $\gcd(k, n)$ is a positive divisor of n and by the previous proposition each divisor d of n is equal to $\gcd(k, n)$ for precisely $\varphi(\frac{n}{d})$ integers $k \in \{1, 2, \dots, n\}$. We deduce that

$$n = \sum_{d \mid n} \varphi\left(\frac{n}{d}\right).$$

When d runs over all positive divisors of n , so does $\frac{n}{d}$. Thus

$$\sum_{d|n} \varphi\left(\frac{n}{d}\right) = \sum_{d|n} \varphi(d)$$

and the result follows. \square

Example 4.113. Prove that

$$\sum_{d=1}^n \varphi(d) \left\lfloor \frac{n}{d} \right\rfloor = \frac{n(n+1)}{2}.$$

Proof. Since $\left\lfloor \frac{n}{d} \right\rfloor$ is the number of multiples of d in $\{1, 2, \dots, n\}$ we obtain

$$\sum_{d=1}^n \varphi(d) \left\lfloor \frac{n}{d} \right\rfloor = \sum_{d=1}^n \varphi(d) \sum_{\substack{1 \leq k \leq n \\ d|k}} 1 = \sum_{k=1}^n \sum_{\substack{1 \leq d \leq n \\ d|k}} \varphi(d).$$

By Gauss' theorem

$$\sum_{\substack{1 \leq d \leq n \\ d|k}} \varphi(d) = k$$

for all $1 \leq k \leq n$ and the result follows from

$$1 + 2 + \dots + n = \frac{n(n+1)}{2}. \quad \square$$

Example 4.114. (AMM E 3106) For $n > 1$ let $S(n)$ be the set of positive integers k for which the fractional part of $\frac{n}{k}$ is at least $\frac{1}{2}$. Prove that

$$\sum_{k \in S(n)} \varphi(k) = n^2.$$

Proof. The key observation is that for any $k \geq 1$ we have $\left\lfloor \frac{2n}{k} \right\rfloor - 2 \left\lfloor \frac{n}{k} \right\rfloor \in \{0, 1\}$ and $\left\lfloor \frac{2n}{k} \right\rfloor - 2 \left\lfloor \frac{n}{k} \right\rfloor = 1$ if and only if $k \in S(n)$. This follows directly from the identity

$$\lfloor 2x \rfloor - 2 \lfloor x \rfloor = \lfloor 2\{x\} \rfloor,$$

where $\{x\} = x - \lfloor x \rfloor$ is the fractional part of x . We deduce that

$$\sum_{k \in S(n)} \varphi(k) = \sum_{k=1}^{2n} \varphi(k) \left(\left\lfloor \frac{2n}{k} \right\rfloor - 2 \left\lfloor \frac{n}{k} \right\rfloor \right) = \sum_{k=1}^{2n} \varphi(k) \left\lfloor \frac{2n}{k} \right\rfloor - 2 \sum_{k=1}^{2n} \varphi(k) \left\lfloor \frac{n}{k} \right\rfloor.$$

Since $\lfloor \frac{n}{k} \rfloor = 0$ for $k \in \{n+1, \dots, 2n\}$ and since (by the previous example)

$$\sum_{k=1}^N \varphi(k) \left\lfloor \frac{N}{k} \right\rfloor = \frac{N(N+1)}{2},$$

we deduce that

$$\sum_{k \in S(n)} \varphi(k) = \frac{2n(2n+1)}{2} - 2 \frac{n(n+1)}{2} = n^2. \quad \square$$

Example 4.115. (China TST 2014) If $n > 1$, let $f(n)$ be the number of ways of factoring n into a product of integers greater than 1 (the order of factors does not count). For instance $f(12) = 4$ since the corresponding factorizations are $12, 2 \cdot 6, 3 \cdot 4, 2 \cdot 2 \cdot 3$. Prove that for any $n > 1$ and any prime divisor p of n we have $f(n) \leq \frac{n}{p}$.

Proof. We prove this by strong induction, the base case being clear. Assume now that it holds for all numbers less than n and let us prove it for n . Let p be the largest prime divisor of n . Clearly it suffices to prove that $f(n) \leq \frac{n}{p}$. If $n = x_1 x_2 \dots x_k$ is a factorization of n into a product of integers greater than 1, then some x_i is divisible by p , say $x_i = pd$ for some d . Then $d \mid \frac{n}{p}$ and $\frac{n}{pd} = x_1 \dots x_{i-1} x_{i+1} \dots x_k$ is a factorization of $\frac{n}{pd}$ into a product of integers greater than 1. Since there are at most $f\left(\frac{n}{pd}\right)$ such factorizations, we obtain

$$f(n) \leq \sum_{d \mid \frac{n}{p}} f\left(\frac{n}{pd}\right).$$

By the inductive hypothesis for each $k < n$ we have $f(k) \leq \frac{k}{P(k)}$, where $P(k)$ is the largest prime factor of k . We have $\frac{k}{P(k)} \leq \varphi(k)$, since

$$\frac{\varphi(k)}{k} = \prod_{p \mid k} \left(1 - \frac{1}{p}\right) \geq \prod_{i=2}^{P(k)} \left(1 - \frac{1}{i}\right) = \frac{1}{P(k)}.$$

Thus

$$f(n) \leq \sum_{d|\frac{n}{p}} f\left(\frac{n}{pd}\right) \leq \sum_{d|\frac{n}{p}} \varphi\left(\frac{n}{pd}\right) = \frac{n}{p},$$

where the last equality follows from Gauss' theorem.

This finishes the proof. \square

The previous results are of fundamental importance, and it is crucial to get familiar with them in order to understand some of the deeper theorems to come. We will therefore illustrate these theoretical results with quite a few examples.

Example 4.116. Find all positive integers n for which $\varphi(2^{2^n} - 1) = \varphi(2^{2^n})$.

Proof. Let $F_k = 2^{2^k} + 1$ be the k th Fermat number. Since the Fermat numbers are pairwise relatively prime and

$$2^{2^n} - 1 = F_0 \cdot F_1 \cdot \dots \cdot F_{n-1},$$

we can write the equation as

$$\prod_{i=0}^{n-1} \varphi(F_i) = 2^{2^n-1},$$

thanks to the multiplicative character of Euler's function. If $n \geq 6$, we deduce that $\varphi(F_5)$ is a power of 2. Since $641 \mid F_5$ (see example 2.12), 640 divides $\varphi(F_5)$ and $\varphi(F_5)$ is not a power of 2. Thus any solution satisfies $n \leq 5$. Conversely, if $n \leq 5$, then F_i is a prime for $i \leq n-1$, hence

$$\prod_{i=0}^{n-1} \varphi(F_i) = \prod_{i=0}^{n-1} (F_i - 1) = 2^{1+2+\dots+2^{n-1}} = 2^{2^n-1}.$$

Thus the answer is $n = 1, 2, 3, 4, 5$. \square

Example 4.117. Prove that for all integers $n > 1$ one can find integers x for which $\varphi(x) = n!$.

Proof. We will choose x having the same set of prime divisors as $n!$. In this case the equation becomes

$$x \cdot \prod_{p|n!} \left(1 - \frac{1}{p}\right) = n!$$

and is equivalent to

$$x = \frac{n!}{\prod_{p|n!} \left(1 - \frac{1}{p}\right)} = \prod_{p|n!} p \cdot \frac{n!}{\prod_{p|n!} (p-1)}.$$

It is apparent that this x really has the same set of prime factors as $n!$, hence it is a solution of the problem. \square

Example 4.118. (USA TST 2015) Let $\varphi(n)$ denote the number of positive integers less than n that are relatively prime to n . Prove that there exists a positive integer m for which the equation $\varphi(n) = m$ has at least 2015 solutions in n .

Proof. Let p_1, p_2, \dots be the increasing sequence of primes and fix a positive integer k . Define

$$N = p_1 p_2 \dots p_k \quad \text{and} \quad x_i = N \left(1 - \frac{1}{p_i}\right), \quad 1 \leq i \leq k.$$

We claim that $\varphi(x_i) = \varphi(N)$ for $1 \leq i \leq k$ (thus taking $k = 2015$ solves the problem). Indeed, since all prime factors of $p_i - 1$ are among p_1, \dots, p_{i-1} , it follows that the prime factors of x_i are exactly $p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_k$ (each appearing with a certain multiplicity). Thus

$$\begin{aligned} \frac{\varphi(x_i)}{x_i} &= \prod_{j=1}^{i-1} \left(1 - \frac{1}{p_j}\right) \cdot \prod_{j=i+1}^k \left(1 - \frac{1}{p_j}\right) \\ &= \prod_{j=1}^k \left(1 - \frac{1}{p_j}\right) \cdot \frac{1}{\left(1 - \frac{1}{p_i}\right)} = \frac{\varphi(N)}{N} \cdot \frac{1}{\left(1 - \frac{1}{p_i}\right)} = \frac{\varphi(N)}{x_i}. \end{aligned}$$

Thus $\varphi(x_i) = \varphi(N)$, as needed. \square

Remark 4.119. A theorem of Pillai shows that $\lim_{n \rightarrow \infty} \frac{f(n)}{n} = 0$, where $f(n)$ is the number of $x \in \{1, 2, \dots, n\}$ that are also in the image of Euler's totient function. This immediately implies the result of the previous example, but the proof of Pillai's theorem requires some delicate estimates for primes, which are totally avoided by the beautiful argument (due to Schinzel) explained in the previous proof.

Example 4.120. Prove that for all $n > 1$ we have:

- a) $\sigma(n) < n(1 + \log n)$;
- b) $n^2 > \sigma(n) \cdot \varphi(n) > \frac{n^2}{2}$;
- c) $\varphi(n) > \frac{n}{4 \log n}$.

Proof. Part c) follows directly by combining parts a) and b).

- a) When d runs over the positive divisors of n , so does $\frac{n}{d}$, hence

$$\sigma(n) = \sum_{d|n} d = \sum_{d|n} \frac{n}{d} = n \sum_{d|n} \frac{1}{d}.$$

Using the inequality

$$\sum_{d=1}^n \frac{1}{d} < 1 + \log n$$

we obtain

$$\frac{\sigma(n)}{n} = \sum_{d|n} \frac{1}{d} \leq \sum_{d=1}^n \frac{1}{d} < 1 + \log n.$$

- b) If $n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$ is the prime factorization of n , with $p_1 < p_2 < \dots < p_k$ then

$$\frac{\sigma(n)}{n} = \prod_{i=1}^k \left(1 + \frac{1}{p_i} + \dots + \frac{1}{p_i^{\alpha_i}} \right) < \prod_{i=1}^k \frac{1}{1 - \frac{1}{p_i}} = \frac{n}{\varphi(n)},$$

thus $\sigma(n)\varphi(n) < n^2$. Next,

$$\sigma(n) \cdot \varphi(n) \geq n \cdot \prod_{i=1}^k \left(1 + \frac{1}{p_i} \right) \cdot n \cdot \prod_{i=1}^k \left(1 - \frac{1}{p_i} \right),$$

hence it suffices to prove that

$$\prod_{i=1}^k \left(1 - \frac{1}{p_i^2}\right) > \frac{1}{2}.$$

This follows from Bernoulli's inequality³ and the inequality

$$\sum_{i=1}^k \frac{1}{p_i^2} < \frac{1}{2}$$

that has already been seen (see example 4.77). \square

Remark 4.121. With a lot more work, one can prove the existence of a constant $c > 0$ (which can be made explicit) such that for all $n > 2$ we have

$$\varphi(n) > c \cdot \frac{n}{\log \log n}.$$

Example 4.122. (Romania TST 2014) Let n be a positive integer and let A_n (respectively B_n) be the set of integers $k \in \{1, 2, \dots, n\}$ such that $\gcd(k, n)$ has an even (respectively odd) number of prime factors (without counting multiplicities). Prove that $|A_n| = |B_n|$ for n even and $|A_n| > |B_n|$ for n odd. Note: 1 has 0 prime factors.

Proof. Let $\omega(k)$ be the number of distinct prime factors of k . Then clearly $\omega(xy) = \omega(x) + \omega(y)$ when x, y are relatively prime, thus $x \mapsto (-1)^{\omega(x)}$ is multiplicative. Next, by definition

$$|A_n| - |B_n| = \sum_{k=1}^n (-1)^{\omega(\gcd(n, k))}.$$

For each divisor $d \mid n$ there are precisely $\varphi\left(\frac{n}{d}\right)$ integers $k \in \{1, 2, \dots, n\}$ such that $\gcd(k, n) = d$. Thus

$$|A_n| - |B_n| = \sum_{d \mid n} (-1)^{\omega(d)} \varphi\left(\frac{n}{d}\right).$$

³This says that $(1 - x_1)(1 - x_2) \dots (1 - x_n) \geq 1 - (x_1 + \dots + x_n)$ for all $x_1, \dots, x_n \in [0, 1]$. The proof is a simple induction on n , left to the reader.

In other words, the map $n \mapsto |A_n| - |B_n|$ is the convolution product of two multiplicative functions $n \mapsto (-1)^{\omega(n)}$ and $n \mapsto \varphi(n)$. Thus by theorem 4.99 the function $n \mapsto |A_n| - |B_n|$ is itself multiplicative, and so it suffices to know its values on prime powers. If $n = p^k$ with $k \geq 1$ and p a prime, it is clear that

$$\begin{aligned} |A_n| - |B_n| &= \sum_{j=1}^{p^k} (-1)^{\omega(\gcd(p^k, j))} = \sum_{p|j} (-1) + \sum_{\gcd(j, p)=1} 1 \\ &= -p^{k-1} + p^k - p^{k-1} = p^{k-1}(p-2). \end{aligned}$$

We conclude that for all n we have

$$|A_n| - |B_n| = n \prod_{p|n} \left(1 - \frac{2}{p}\right)$$

and the result follows. \square

4.4.4 The Möbius function and its applications

In this section we discuss in more detail some basic properties of the Möbius function μ . Recall that it is defined by $\mu(1) = 1$, $\mu(n) = 0$ whenever n is not squarefree (i.e. n is not a multiple of p^2 for any prime p) and $\mu(n) = (-1)^{\omega(n)}$ when n is squarefree. Its key property is the following relation (the reader should be careful, the relation below only holds for $n > 1$, not for $n = 1$).

Proposition 4.123. *We have $\sum_{d|n} \mu(d) = 0$ for $n > 1$.*

Proof. Let $n = p_1^{\alpha_1} \cdots p_m^{\alpha_m}$ be the prime factorization of n . In the sum $\sum_{d|n} \mu(d)$, the only d 's giving nonzero contributions are 1, the prime factors of n , the products of two distinct prime factors of n , ..., up to $p_1 \cdots p_m$. Since there are $\binom{m}{j}$ products of j distinct prime divisors of n , and each such product has contribution $(-1)^j$, we obtain

$$\sum_{d|n} \mu(d) = 1 - \binom{m}{1} + \binom{m}{2} - \cdots = (1-1)^m = 0,$$

using the binomial theorem. The result follows. \square

An important consequence of the previous proposition is the famous Möbius inversion formula:

Theorem 4.124. (*Möbius inversion formula*) If $f(n) = \sum_{d|n} g(d)$ for all n , then

$$g(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) f(d)$$

for all n .

Proof. We compute

$$\sum_{d|n} \mu\left(\frac{n}{d}\right) f(d) = \sum_{d|n} \mu\left(\frac{n}{d}\right) \cdot \sum_{e|d} g(e) = \sum_{e|n} g(e) \cdot \sum_{e|d|n} \mu\left(\frac{n}{d}\right).$$

On the other hand, writing $d = ex$, we have $x \mid \frac{n}{e}$ and $\frac{n}{d} = \frac{e}{x}$, thus by proposition 4.123 we have

$$\sum_{e|d|n} \mu\left(\frac{n}{d}\right) = \sum_{x|\frac{n}{e}} \mu\left(\frac{e}{x}\right) = 0$$

unless $e = n$, in which case the sum is equal to 1. The result follows. \square

Remark 4.125. 1) There is also a multiplicative version of the Möbius inversion formula (proved in exactly the same way): if

$$f(n) = \prod_{d|n} g(d)$$

for all n , then

$$g(n) = \prod_{d|n} f(d)^{\mu\left(\frac{n}{d}\right)}.$$

2) The same argument shows that if f, g are arithmetic functions related by

$$g(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) f(d)$$

for all n , then

$$f(n) = \sum_{d|n} g(d)$$

for all n . In other words, the converse of the previous theorem holds. Indeed, we have

$$\sum_{d|n} g(d) = \sum_{d|n} \sum_{e|d} \mu\left(\frac{d}{e}\right) f(e) = \sum_{e|n} f(e) \sum_{e|d|n} \mu\left(\frac{d}{e}\right)$$

and

$$\sum_{e|d|n} \mu\left(\frac{d}{e}\right) = \sum_{x|\frac{n}{e}} \mu(x)$$

equals 1 if $e = n$ and 0 otherwise.

3) Sometimes it can be useful to consider functions f, g which are only defined on the set of positive divisors of a fixed number $N > 1$. If they satisfy

$$f(n) = \sum_{d|n} g(d)$$

for any $n | N$, then we can still deduce (using the same arguments as above) that

$$g(n) = \sum_{d|n} g(d) \mu\left(\frac{n}{d}\right)$$

for any $n | N$. We leave the details as an exercise to the reader.

Let us apply now the previous results to Euler's function φ . Consider Gauss' identity (see theorem 4.112)

$$n = \sum_{d|n} \varphi(d)$$

and apply the Möbius inversion formula to it. We obtain

$$\begin{aligned} \varphi(n) &= \sum_{d|n} \mu\left(\frac{n}{d}\right) d = \sum_{d|n} \mu(d) \frac{n}{d} = n \sum_{d|n} \frac{\mu(d)}{d} \\ &= n \left(1 - \sum_{p|n} \frac{1}{p} + \sum_{p < q | n} \frac{1}{pq} - \dots \right) = n \prod_{p|n} \left(1 - \frac{1}{p} \right). \end{aligned}$$

In other words, we recover the formula

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

that we obtained in the previous section. Conversely, using the previous formula and the Möbius inversion formula we recover Gauss' theorem. Thus Gauss' theorem and the explicit formula for $\varphi(n)$ are actually equivalent!

The following beautiful result is a nice illustration of the Möbius inversion formula.

Example 4.126. Let $(a_n)_{n \geq 1}$ be a sequence of positive integers such that

$$\gcd(a_m, a_n) = a_{\gcd(m, n)}$$

for all positive integers m, n . Prove that there exists a sequence of positive integers $(b_n)_{n \geq 1}$ such that for all $n \geq 1$

$$a_n = \prod_{d|n} b_d.$$

Proof. By the multiplicative version of the Möbius inversion formula we have

$$b_n = \prod_{d|n} a_{\frac{n}{d}}^{\mu(d)}$$

and we need to prove that this is an integer for all n . Letting p_1, \dots, p_d be the (pairwise distinct) prime factors of n , we obtain

$$b_n = \frac{a_n}{\prod_{i=1}^d a_{\frac{n}{p_i}}} \cdot \frac{\prod_{i < j} a_{\frac{n}{p_i p_j}}}{\prod_{i < j < k} a_{\frac{n}{p_i p_j p_k}}} \cdot \dots$$

On the other hand, using the hypothesis of the problem repeatedly yields

$$a_{\frac{n}{p_i p_j}} = \gcd(a_{\frac{n}{p_i}}, a_{\frac{n}{p_j}}), \quad a_{\frac{n}{p_i p_j p_k}} = \gcd(a_{\frac{n}{p_i}}, a_{\frac{n}{p_j}}, a_{\frac{n}{p_k}}), \dots$$

Letting $x_i = a_{\frac{n}{p_i}}$ for $1 \leq i \leq d$, we deduce that

$$b_n = \frac{a_n}{\prod_{i=1}^d x_i} \cdot \frac{\prod_{i < j} \gcd(x_i, x_j)}{\prod_{i < j < k} \gcd(x_i, x_j, x_k)} \cdot \dots$$

The lemma 4.127 below yields therefore

$$b_n = \frac{a_n}{\text{lcm}(a_{\frac{n}{p_1}}, \dots, a_{\frac{n}{p_d}})},$$

an expression which makes it clear that b_n is an integer. \square

Lemma 4.127. *For any positive integers x_1, \dots, x_d we have*

$$\frac{\prod_{i=1}^d x_i}{\prod_{i < j} \gcd(x_i, x_j)} \cdot \frac{\prod_{i < j < k} \gcd(x_i, x_j, x_k)}{\prod_{i < j < k < l} \gcd(x_i, x_j, x_k, x_l)} \cdot \dots = \text{lcm}(x_1, \dots, x_d).$$

Proof. The result is clear for $d = 2$ and assuming that it holds for d , we obtain

$$\text{lcm}(x_1, \dots, x_{d+1}) = \text{lcm}(\text{lcm}(x_1, \dots, x_d), x_{d+1}) = \frac{x_{d+1} \cdot \text{lcm}(x_1, \dots, x_d)}{\gcd(x_{d+1}, \text{lcm}(x_1, \dots, x_d))}.$$

Inserting the value of $\text{lcm}(x_1, \dots, x_d)$ given by the inductive hypothesis in the previous expression yields the desired result (after some tedious but simple algebraic manipulations left to the reader). \square

4.4.5 Application to squarefree numbers

We want now to use the Möbius function in order to study the distribution of squarefree numbers. We warn the reader that the remainder of this section is rather technical, so he can freely skip what follows for a first lecture.

Let $Q(n)$ be the number of squarefree numbers between 1 and n and let P be the set of primes $p \leq \sqrt{n}$. Define for each $p \in P$ the set

$$A_p = \{x \in \{1, 2, \dots, n\} \mid p^2 \mid x\}.$$

Then the set of squarefree numbers between 1 and n is precisely the set $\{1, 2, \dots, n\} \setminus \bigcup_{p \in P} A_p$, thus using the inclusion-exclusion principle we obtain

$$Q(n) = n - \sum_{p \in P} |A_p| + \sum_{p < q \in P} |A_p \cap A_q| + \dots$$

On the other hand, since there are $\left\lfloor \frac{n}{k} \right\rfloor$ multiples of k between 1 and n , we deduce that

$$|A_{p_1} \cap A_{p_2} \cap \dots \cap A_{p_k}| = \left\lfloor \frac{n}{p_1^2 \dots p_k^2} \right\rfloor$$

for $p_1 < \dots < p_k \in P$. We conclude that

$$Q(n) = n - \sum_{p \in P} \left\lfloor \frac{n}{p^2} \right\rfloor + \sum_{p < q \in P} \left\lfloor \frac{n}{p^2 q^2} \right\rfloor - \dots = \sum_{k \leq \sqrt{n}} \mu(k) \left\lfloor \frac{n}{k^2} \right\rfloor,$$

in other words we have just proved the

Proposition 4.128. *The number of squarefree numbers between 1 and n is given by*

$$Q(n) = \sum_{k \leq \sqrt{n}} \mu(k) \left\lfloor \frac{n}{k^2} \right\rfloor.$$

Noting that $\mu(k)$ takes only the values $-1, 0, 1$ and the distance between $\left\lfloor \frac{n}{k^2} \right\rfloor$ and $\frac{n}{k^2}$ is at most 1, we obtain

$$\left| Q(n) - n \sum_{k \leq \sqrt{n}} \frac{\mu(k)}{k^2} \right| \leq \sqrt{n}.$$

This shows that in order to estimate $Q(n)$ we need to estimate $\sum_{k \leq \sqrt{n}} \frac{\mu(k)}{k^2}$. The key ingredient is the following remarkable identity, which looks very similar to Euler's famous identity

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots = \frac{\pi^2}{6}.$$

Actually the proof will show that the next theorem is equivalent to this identity.

Theorem 4.129. *We have*

$$\sum_{k=1}^{\infty} \frac{\mu(k)}{k^2} = \frac{6}{\pi^2}.$$

Proof. Using Euler's identity, it suffices to prove that

$$\sum_{j \geq 1} \frac{1}{j^2} \cdot \sum_{k \geq 1} \frac{\mu(k)}{k^2} = 1.$$

Expanding, the left-hand side equals

$$\sum_{j, k \geq 1} \frac{\mu(k)}{(jk)^2} = \sum_{n \geq 1} \sum_{jk=n} \frac{\mu(k)}{n^2} = \sum_{n \geq 1} \frac{1}{n^2} \sum_{k|n} \mu(k) = 1,$$

the last identity being a consequence of the fact that $\sum_{d|k} \mu(d)$ equals 0 for $k > 1$ and 1 for $k = 1$. \square

We are now in good shape for proving the following beautiful result:

Theorem 4.130. *The number $Q(n)$ of squarefree numbers between 1 and $n > 1$ satisfies*

$$\left| Q(n) - \frac{6}{\pi^2} n \right| \leq 3\sqrt{n}.$$

Proof. Using the previously established inequality

$$\left| Q(n) - n \sum_{k \leq \sqrt{n}} \frac{\mu(k)}{k^2} \right| \leq \sqrt{n}$$

as well as the result of the previous theorem we reduce the proof to the inequality

$$\left| \sum_{k > \sqrt{n}} \frac{\mu(k)}{k^2} \right| \leq \frac{2}{\sqrt{n}}.$$

Clearly, it suffices to prove that

$$\sum_{k > \sqrt{n}} \frac{1}{k^2} \leq \frac{2}{\sqrt{n}}.$$

Letting $N = \lfloor \sqrt{n} \rfloor$, we obtain

$$\sum_{k > \sqrt{n}} \frac{1}{k^2} = \sum_{k \geq N+1} \frac{1}{k^2} < \sum_{k \geq N+1} \frac{1}{k(k-1)} = \sum_{k \geq N+1} \left(\frac{1}{k-1} - \frac{1}{k} \right) = \frac{1}{N}.$$

Since $\frac{1}{N} \leq \frac{2}{\sqrt{n}}$, we are done. \square

Example 4.131. Prove that any $n > 1000$ can be written as the sum of two squarefree numbers.

Proof. We claim that

$$Q(n-1) > \frac{n-1}{2}.$$

Using the previous theorem it suffices to prove that

$$\frac{6}{\pi^2}(n-1) > \frac{n-1}{2} + 3\sqrt{n-1},$$

which easily follows from $n > 1000$ and $\frac{6}{\pi^2} > 0.6$. Consider now the set A of all squarefree numbers between 1 and $n-1$ and the set B of all numbers of the form $n-x$, with $x \in A$. Then A and B are subsets of $\{1, 2, \dots, n-1\}$, each with more than $\frac{n-1}{2}$ elements. Thus their intersection is nonempty and so we can find $x \in A$ such that $n-x \in A$. But then $n = x + (n-x)$ is the sum of two squarefree numbers. \square

Remark 4.132. Explicit computations show that any $n > 1$ is a sum of two squarefree numbers.

Example 4.133. Prove that for infinitely many integers $n > 1$ all numbers $n, n+1, n+2$ are squarefree.

Proof. Suppose that there is N such that for all $n \geq N$ at least one of the numbers $n, n+1, n+2$ is not squarefree. Then for each $k \geq N$ at least two of the numbers $4k, 4k+1, 4k+2, 4k+3$ are not squarefree. Dividing the numbers $4N, 4N+1, \dots, 4n-1$ into groups of 4 consecutive numbers, we deduce that

$$Q(4n) - Q(4N) \leq 2n$$

for all $n \geq N$. This is impossible, since by the previous theorem

$$\lim_{n \rightarrow \infty} \frac{Q(4n)}{2n} = \frac{12}{\pi^2} > 1. \quad \square$$

Example 4.134. Let a_1, \dots, a_d and b_1, \dots, b_d be positive integers. Prove that if there is an integer n such that $a_1n+b_1, \dots, a_dn+b_d$ are all squarefree numbers, then there are infinitely many such $n > 0$.

Proof. Fix an integer n_0 such that $a_i n_0 + b_i$ are all squarefree numbers, and let C be a large number (we will make a final choice later on), such that any prime factor of $\prod_{i=1}^d a_i(a_i n_0 + b_i)$ is smaller than C . Let P be the product of all primes not exceeding C . We will prove that

$$x_i(k) := a_i(n_0 + kP^2) + b_i = a_i n_0 + b_i + kP^2 a_i$$

are all squarefree numbers for infinitely many positive integers k , which is enough to conclude.

Fix a large integer $N > C$ and consider those $x_i(k)$ with $1 \leq i \leq d$ and $1 \leq k \leq N$. Note that $x_i(k)$ is not a multiple of p^2 for any prime $p \leq C$ (as otherwise $a_i n_0 + b_i$ would be a multiple of p^2). Assume that for some $i \leq d$ the number $x_i(k)$ is not squarefree, thus there is a prime $p > C$ such that $p^2 \mid x_i(k)$. Then (if C is large enough)

$$p^2 \leq x_i(k) < 2kP^2 a_i \leq 2NP^2 a_i,$$

thus $p < \sqrt{2Na_i}P \leq \sqrt{2MNP}$, where $M = \max(a_1, \dots, a_d)$. Moreover, since p does not divide Pa_i , the solutions of the congruence $x_i(k) \equiv 0 \pmod{p^2}$ considered as a linear congruence in k are all congruent modulo p^2 , so there are at most $1 + \frac{N}{p^2}$ such solutions. Since there are less than $\sqrt{2NMP}$ primes below $\sqrt{2NMP}$, we deduce that $x_i(k)$ is not squarefree for at most

$$\sqrt{2NMP} + N \sum_{p>C} \frac{1}{p^2} < \sqrt{2NMP} + N \sum_{k>C} \frac{1}{k(k-1)} < \sqrt{2NMP} + \frac{N}{C}$$

values of $k \in [1, N]$. Therefore all numbers $x_1(k), x_2(k), \dots, x_d(k)$ are squarefree for at least

$$N \left(1 - \frac{d}{C}\right) - dP\sqrt{2M} \cdot \sqrt{N}$$

values of $k \in [1, N]$. Since the last quantity tends to ∞ as $N \rightarrow \infty$ (fixing once and for all $C > d$ large enough), the result follows. \square

Example 4.135. (IMC 2013) Is there an infinite set of positive integers A such that for all distinct elements $a, b \in A$ the number $a + b$ is squarefree?

Proof. We will construct inductively an infinite increasing sequence $a_1 < a_2 < \dots$ such that $a_1 = 1, a_2 = 2$ and $a_i + a_j$ is square free whenever $i \neq j$. Assume that a_1, \dots, a_k have already been constructed, we will try to construct a_{k+1} so that $a_{k+1} + a_i$ are square free for $1 \leq i \leq k$. Consider two auxiliary big numbers r, N and let us look for a_{k+1} of the form $1 + r!n$ for some $n \in \{1, 2, \dots, N\}$. We will choose $r > k + \max_{1 \leq i \leq k} (1 + a_i)^2$ to ensure that $1 + r!n + a_i$ is of the form $(1 + a_i)(1 + y(1 + a_i))$ for some $y \geq 1$. Thus if $p^2 \mid 1 + r!n + a_i$ for some $1 \leq i \leq k$ and some prime p , then necessarily $p > r$ (if $p \leq r$ then $p \mid 1 + a_i$ and then necessarily $p^2 \mid 1 + a_i$, contradicting the fact that $a_i + a_1 = a_i + 1$ is square free, by the inductive hypothesis if $i > 1$ and by the choice of a_1 for $i = 1$). Moreover, $p^2 \leq 1 + r!n + a_i < r!(N + 1)$. There are at most $\frac{N}{p^2} + 1$ values of $n \in \{1, 2, \dots, N\}$ for which $p^2 \mid 1 + r!n + a_i$, thus in total there are at most

$$S = k \cdot \sum_{r < p < \sqrt{r!(N+1)}} \left(\frac{N}{p^2} + 1 \right)$$

numbers $n \in \{1, 2, \dots, N\}$ for which $1 + r!n + a_i$ is not squarefree for some $1 \leq i \leq k$. Note that

$$\begin{aligned} S &< k \left(\sqrt{r!(N+1)} + N \sum_{j>r} \frac{1}{j^2} \right) < k\sqrt{r!(N+1)} + kN \sum_{j>r} \left(\frac{1}{j-1} - \frac{1}{j} \right) \\ &< k\sqrt{r!(N+1)} + \frac{k}{r}N \end{aligned}$$

and the last expression is less than $N - 1$ for N big enough since $k < r$. Thus for N big enough (and with any fixed choice of $r > k + \max_{1 \leq i \leq k} (1 + a_i)^2$) we can choose $a_{k+1} = 1 + r!n$ for some $n \in \{1, 2, \dots, N\}$ to make $a_{k+1} + a_i$ squarefree for $1 \leq i \leq k$, finishing the inductive step. \square

Example 4.136. (Brazil 2015) If $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$ is the prime factorization of $n > 1$, let

$$f(n) = \alpha_1 p_1^{\alpha_1 - 1} \alpha_2 p_2^{\alpha_2 - 1} \dots \alpha_k p_k^{\alpha_k - 1}.$$

Prove that $f(n) = f(n - 1) + 1$ for infinitely many n .

Proof. Note that $f(n) = 1$ whenever n is squarefree and that f is clearly multiplicative. Let

$$a = 27, \quad b = 169, \quad x = 482, \quad y = 77.$$

Then x, y are squarefree, $ax = by + 1$, $\gcd(a, x) = \gcd(b, y) = 1$ and $f(a) = f(b) + 1$. By example 4.134 the numbers $ab^2n + x$ and $a^2bn + y$ are simultaneously squarefree for infinitely many $n \geq 1$, thus for such n we have

$$\begin{aligned} f(a^2b^2n + ax) &= f(a(ab^2n + x)) = f(a) = 1 + f(b) \\ &= 1 + f(b(a^2bn + y)) = 1 + f(a^2b^2n + ax - 1). \end{aligned}$$

Thus $f(m) = f(m - 1) + 1$ for $m = a^2b^2n + ax$ and n as above. The result follows. \square

4.5 Problems for practice

Composite numbers

1. Prove that if a is an integer greater than 1 and if $n > 1$ is not a power of 2, then $a^n + 1$ is composite.
2. (St. Petersburg 2004) Prove that for any integer a there exist infinitely many positive integers n such that $a^{2^n} + 2^n$ is composite.
3. Find all positive integers n for which at least one of the numbers $n^n + 1$ and $(2n)^{2n} + 1$ is composite.
4. For which positive integers n the numbers $2^n + 3$ and $2^n + 5$ are both primes?
5. (St. Petersburg 1996) Integers a, b, c have the property that the roots of the polynomial $X^3 + aX^2 + bX + c$ are pairwise relatively prime and distinct positive integers. Prove that if the polynomial $aX^2 + bX + c$ has a positive integer root, then $|a|$ is composite.

6. (Vojtech Jarnik Competition 2009) Prove that if $k > 2$ then $2^{2^k-1} - 2^k - 1$ is composite.
7. A positive integer which is congruent 1 modulo 4 has two different representations as a sum of two squares. Prove that this number is composite.
8. (Moscow Olympiad) Is there an 1997-digit composite number such that if any three of its consecutive digits are replaced by any other triplet of digits then the resulting number is composite?
9. (AMM 10947) Prove that $\frac{5^{5n}-1}{5^n-1}$ is composite for all $n \geq 1$.

The fundamental theorem of arithmetic

10. Let $n > 1$ be an integer. Prove that the equation

$$(x+1)(x+2)\dots(x+n) = y^n$$

has no solution in positive integers.

11. Let n be a positive integer. Prove that if n divides $\binom{n}{k}$ for all $1 \leq k \leq n-1$, then n is prime.
12. (USAMTS 2009) Find a positive integer n such that all prime factors of

$$\frac{(n+1)(n+2)\dots(n+500)}{500!}$$

are greater than 500.

13. (Russia 1999) Prove that any positive integer is the difference of two positive integers with the same number of prime factors (without counting multiplicities).
14. (Saint Petersburg) An infinite sequence $(a_n)_{n \geq 1}$ of composite numbers satisfies

$$a_{n+1} = a_n - p_n + \frac{a_n}{p_n}$$

for all n , where p_n is the smallest prime factor of a_n . If all terms of the sequence are multiples of 37, what are the possible values of a_1 ?

15. Prove that there are infinitely many pairs (a, b) of distinct positive integers a, b such that a and b have the same prime divisors, and $a + 1$ and $b + 1$ also have the same prime divisors.
16. Let a, b, c, d, e, f be positive integers such that $abc = def$. Prove that $a(b^2 + c^2) + d(e^2 + f^2)$ is composite.
17. (Kvant M 1762) Is there a positive integer n with 2013 prime divisors such that n divides $2^n + 1$?
18. (Poland 2000) Let p_1 and p_2 be prime numbers and for $n \geq 3$ let p_n be the greatest prime factor of $p_{n-1} + p_{n-2} + 2000$. Prove that the sequence $(p_n)_{n \geq 1}$ is bounded.
19. (Italy 2011) Find all primes p for which $p^2 - p - 1$ is the cube of an integer.
20. (Kvant M 2145) Let $x > 2, y > 1$ be integers such that $x^y + 1$ is a perfect square. Prove that x has at least 3 different prime divisors.
21. (Russia 2010) Prove that for any $n > 1$ there are n consecutive positive integers whose product is divisible by all primes not exceeding $2n + 1$, and not divisible by any other prime.
22. (Iran 2015) Prove that infinitely many positive integers n cannot be written as the sum of two positive integers all of whose prime factors are less than 1394.
23. (China 2007) Let $n > 1$ be an integer. Prove that $2n - 1$ is a prime number if and only if for any n pairwise distinct positive integers a_1, a_2, \dots, a_n there exist $i, j \in \{1, 2, \dots, n\}$ such that

$$\frac{a_i + a_j}{\gcd(a_i, a_j)} \geq 2n - 1$$

24. (Tournament of the Towns 2009) Initially the number 6 is written on a blackboard. At the n th step, one replaces the number d written on the

blackboard with $d + \gcd(d, n)$. Prove that at each step the number on the blackboard increases either by 1 or by a prime number.

Infinitude of primes

25. (Komal) Is it possible to find 2000 positive integers such that none of them is divisible by any of the other numbers but the square of each is divisible by all the others?
26. A positive integer n is called powerful if $p^2 \mid n$ for any prime factor p of n . Prove that there are infinitely many pairs of consecutive powerful numbers.
27. Let p_n be the largest prime not exceeding n and let q_n be the smallest prime larger than n . Prove that for all $n > 1$ we have

$$\sum_{k=2}^n \frac{1}{p_k q_k} < \frac{1}{2}.$$

28. (Russia 2010) Are there infinitely many positive integers which cannot be expressed as $\frac{x^2-1}{y^2-1}$, with x, y integers greater than 1?
29. (Baltic Way 2004) Is there an infinite sequence of prime numbers p_1, p_2, \dots such that $|p_{n+1} - 2p_n| = 1$ for each $n \geq 1$?
30. Let a_1, a_2, \dots, a_k be positive real numbers such that for all but finitely many positive integers n we have

$$\gcd(n, \lfloor a_1 n \rfloor + \lfloor a_2 n \rfloor + \dots + \lfloor a_k n \rfloor) > 1.$$

Prove that a_1, \dots, a_k are integers.

31. (IMO Shortlist 2006) We define a sequence a_1, a_2, a_3, \dots by setting

$$a_n = \frac{1}{n} \left(\left\lceil \frac{n}{1} \right\rceil + \left\lceil \frac{n}{2} \right\rceil + \dots + \left\lceil \frac{n}{n} \right\rceil \right)$$

for every positive integer n .

- a) Prove that $a_{n+1} > a_n$ for infinitely many n .
- b) Prove that $a_{n+1} < a_n$ for infinitely many n .
32. (APMO 1994) Find all integers n of the form $a^2 + b^2$ with a, b relatively prime positive integers, such that any prime $p \leq \sqrt{n}$ divides ab .
33. (Iran TST 2009) Find all polynomials f with integer coefficients having the following property: for all primes p and for all integers a, b , if $p \mid ab - 1$, then $p \mid f(a)f(b) - 1$.
34. Prove that there is a positive integer n such that the interval $[n^2, (n+1)^2]$ contains at least 2016 primes.
35. (IMO 1977) Let $n > 2$ be an integer and let V_n be the set of integers of the form $1 + kn$ with $k \geq 1$. A number $m \in V_n$ is called indecomposable if it cannot be written as the product of two elements of V_n . Prove that there is $r \in V_n$ that can be expressed as the product of indecomposable elements of V_n in more than one way (expressions which differ only in order of the elements of V_n will be considered the same).
36. (German TST 2009) The sequence $(a_n)_{n \in \mathbb{N}}$ is defined by $a_1 = 1$ and

$$a_{n+1} = a_n^4 - a_n^3 + 2a_n^2 + 1$$

for all $n \geq 1$. Prove that there are infinitely many primes which do not divide any of the numbers a_1, a_2, \dots

Arithmetic functions

37. Prove that for all $n \geq 1$ we have

$$\sum_{d|n} \sigma(d) = n \cdot \sum_{d|n} \frac{\tau(d)}{d}, \quad n \cdot \sum_{d|n} \frac{\sigma(d)}{d} = \sum_{d|n} d\tau(d).$$

38. a) Let f be a multiplicative function with $f(1) = 1$ (this is equivalent to f being nonzero). Prove that for all $n > 1$ we have

$$\sum_{d|n} f(d)\mu(d) = \prod_{p|n} (1 - f(p)),$$

the product being taken over the prime divisors of n .

- b) Deduce closed formulae for

$$\sum_{d|n} \mu(d)\tau(d), \quad \sum_{d|n} \mu(d)\sigma(d) \quad \text{and} \quad \sum_{d|n} \mu(d)\varphi(d)$$

for $n > 1$.

39. Let f be an arithmetic function such that the function g defined by

$$g(n) = \sum_{d|n} f(d)$$

is multiplicative. Prove that f is multiplicative.

40. a) Let f be an arithmetic function and let g be the arithmetic function defined by

$$g(n) = \sum_{d|n} f(d).$$

For all $n \geq 1$ we have

$$\sum_{k=1}^n g(k) = \sum_{k=1}^n f(k) \left[\frac{n}{k} \right].$$

- b) Prove that the following relations hold for all $n \geq 1$

$$\sum_{k=1}^n \tau(k) = \sum_{k=1}^n \left[\frac{n}{k} \right], \quad \sum_{k=1}^n \sigma(k) = \sum_{k=1}^n k \left[\frac{n}{k} \right].$$

41. Let $f(n)$ be the difference between the number of positive divisors of n of the form $3k+1$ and the number of positive divisors of the form $3k-1$. Prove that f is multiplicative.

42. (AMM 2001) Find all totally multiplicative functions $f : \mathbf{N} \rightarrow \mathbf{C}$ such that the function

$$F(n) = \sum_{k=1}^n f(k)$$

is also totally multiplicative.

43. Find all nonzero totally multiplicative functions $f : \mathbf{N} \rightarrow \mathbf{R}$ such that $f(n+1) \geq f(n)$ for all n .
44. (Erdős) Let $f : \mathbf{N} \rightarrow \mathbf{R}$ be a nonzero multiplicative function such that $f(n+1) \geq f(n)$ for all n . Then there is a nonnegative real number k such that $f(n) = n^k$ for all n .
45. Are there infinitely many $n > 1$ such that $n \mid 2^{\sigma(n)} - 1$?
46. An integer $n > 1$ is called perfect if $\sigma(n) = 2n$. Prove that an even number $n > 1$ is perfect if and only if $n = 2^{p-1}(2^p - 1)$, with $2^p - 1$ prime.
47. Let n be an even positive integer. Prove that $\sigma(\sigma(n)) = 2n$ if and only if there is a prime p such that $2^p - 1$ is a prime and $n = 2^{p-1}$.
48. (Romania TST 2010) Prove that for each positive integer a we have $\sigma(an) < \sigma(an+1)$ for infinitely many positive integers n .
49. (IMO Shortlist 2004) Prove that for infinitely many positive integers a the equation $\tau(an) = n$ has no solutions in positive integers.
50. (IMO) Let $\tau(n)$ be the number of divisors of a positive integer n . Find all positive integers k such that $k = \frac{\tau(n^2)}{\tau(n)}$ for some n .
51. A positive integer a is called highly divisible if it has more divisors than any number less than a . If p is a prime number and $a > 1$ is an integer, we write $v_p(a)$ for the exponent of p in the prime factorization of a . Prove that
- a) There are infinitely many highly divisible numbers.

- b) If a is highly divisible and $p < q$ are primes, then $v_p(a) \geq v_q(a)$.
- c) Let p, q be primes such that $p^k < q$ for some positive integer k . Prove that if a is highly divisible and a multiple of q , then a is a multiple of p^k .
- d) Let p, q be primes and let k be a positive integer such that $p^k > q$. Prove that if p^{2k} divides some highly divisible number a , then q divides a .
- e) (China TST 2012) Let n be a positive integer. Prove that all sufficiently large highly divisible numbers are multiples of n .

52. Let $n > 1$ be an integer. Compute

$$\sum_{d|n} (-1)^{\frac{n}{d}} \varphi(d).$$

53. (IMO 1991) Let $1 = a_1 < a_2 < \dots < a_{\varphi(n)}$ be the totatives of $n > 1$. Prove that $a_1, a_2, \dots, a_{\varphi(n)}$ form an arithmetic progression if and only if n is either 6, a prime number or a power of 2.
54. Let $n \geq 2$. Prove that n is a prime if and only if $\varphi(n) \mid n - 1$ and $n + 1 \mid \sigma(n)$ (recall that $\sigma(n)$ is the sum of the positive divisors of n).
55. Let k be a positive integer. Prove that there is a positive integer n such that $\varphi(n) = \varphi(n + k)$.
56. Prove that for all $n \geq 1$ we have

$$\frac{\varphi(1)}{2^1 - 1} + \frac{\varphi(2)}{2^2 - 1} + \dots + \frac{\varphi(n)}{2^n - 1} < 2.$$

57. a) Prove that there are infinitely many integers $n > 1$ such that

$$\varphi(n) \geq \varphi(k) + \varphi(n - k) \text{ for all } 1 \leq k \leq n - 1.$$

- b) Are there infinitely many $n > 1$ such that $\varphi(n) \leq \varphi(k) + \varphi(n - k)$ for all $1 \leq k \leq n - 1$?

58. (AMM 11544) Prove that for any integer $m > 1$ we have

$$\sum_{k=0}^{m-1} \varphi(2k+1) \left\lfloor \frac{m+k}{2k+1} \right\rfloor = m^2.$$

59. a) Prove that for all $n > 1$ we have

$$2 \sum_{k=1}^n \varphi(k) = 1 + \sum_{k=1}^n \mu(k) \left\lfloor \frac{n}{k} \right\rfloor^2.$$

b) Prove that for all $n > 1$ we have

$$\left| \varphi(1) + \varphi(2) + \dots + \varphi(n) - \frac{3}{\pi^2} n^2 \right| < 2n + n \log n.$$

60. Let $a_1, \dots, a_{\varphi(n)}$ be the totatives of $n > 1$.

a) Prove that for all $m \geq 1$ we have

$$a_1^m + a_2^m + \dots + a_{\varphi(n)}^m = \sum_{d|n} \mu(d) d^m \left(1^m + 2^m + \dots + \left(\frac{n}{d} \right)^m \right).$$

b) Compute $a_1^2 + a_2^2 + \dots + a_{\varphi(n)}^2$.

61. (Serbia 2011) Prove that if $n > 1$ is odd and $\varphi(n)$, $\varphi(n+1)$ are powers of 2, then $n+1$ is a power of 2 or $n=5$.

62. (Komal A 492) Let A be a finite set of positive integers. Prove that

$$\sum_{S \subset A} (-2)^{|S|-1} \gcd(S) > 0,$$

the sum running over all nonempty subsets S of A and $\gcd(S)$ denoting the greatest common divisor of all elements of S .

Chapter 5

Congruences involving prime numbers

This long chapter deals with a series of key theorems concerning congruences modulo prime numbers, such as Fermat's little theorem, Wilson's theorem and Langrange's theorem. These are fundamental results in basic number theory, and it is crucial to become very familiar with them before dealing with more advanced results. Therefore we give many concrete examples illustrating each of these results, as well as lots of applications. The second part of the chapter deals with more advanced topics, such as quadratic residues or congruences modulo powers of primes. Once the first part of the chapter is fully understood, the proofs of these more advanced results (with the exception of the quadratic reciprocity law) become relatively simple and natural.

5.1 Fermat's little theorem

5.1.1 Fermat's little theorem and (pseudo-)primality

We now reach the first fundamental congruence in which prime numbers play a key role: Fermat's little theorem. While both the statement and the proof of this theorem are fairly simple, the result itself is incredibly useful, as it will be clear in the sequel.

Theorem 5.1. (*Fermat's little theorem*) For all primes p and all integers a we have

$$a^p \equiv a \pmod{p}.$$

Equivalently, for all primes p and all integers a relatively prime to p we have

$$a^{p-1} \equiv 1 \pmod{p}.$$

Proof. It is clear that the two statements are equivalent, so it suffices to prove the second one. So let a be an integer relatively prime to p . Then $0, a, 2a, 3a, \dots, (p-1)a$ is a complete residue system modulo p by theorem 3.32, hence

$$a \cdot 2a \cdot \dots \cdot (p-1)a \equiv 1 \cdot 2 \cdot \dots \cdot (p-1) \pmod{p}.$$

This can also be written as

$$(p-1)!(a^{p-1} - 1) \equiv 0 \pmod{p}.$$

Since p is a prime, we have $\gcd(p, (p-1)!) = 1$ and so $a^{p-1} \equiv 1 \pmod{p}$, finishing the proof. \square

We would like to explain a second proof of Fermat's little theorem, which is based on a very useful property of binomial coefficients. The reader will find a whole section devoted to congruences between binomial coefficients later on, thus for now we will stick to the simplest one.

Let us recall the classical identity, valid for all $n \geq k \geq 1$

$$k \cdot \binom{n}{k} = n \cdot \binom{n-1}{k-1},$$

which follows from

$$k \cdot \binom{n}{k} = k \cdot \frac{n!}{k!(n-k)!} = \frac{n!}{(k-1)!(n-k)!} = n \cdot \frac{(n-1)!}{(k-1)!(n-k)!} = n \cdot \binom{n-1}{k-1}.$$

We are now ready to state and prove the most fundamental congruence for binomial coefficients:

Theorem 5.2. If p is a prime and $1 \leq k \leq p-1$, then p divides $\binom{p}{k}$.

Proof. The equality $k \binom{p}{k} = p \binom{p-1}{k-1}$ shows that p divides $k \cdot \binom{p}{k}$ and since $\gcd(k, p) = 1$, we conclude that $p \mid \binom{p}{k}$, as desired. \square

We can now explain the second proof of Fermat's little theorem. By theorem 5.2 and the binomial formula we have

$$(x + y)^p - x^p - y^p = \sum_{k=1}^{p-1} \binom{p}{k} x^{p-k} y^k \equiv 0 \pmod{p},$$

that is

$$(x + y)^p \equiv x^p + y^p \pmod{p} \quad (1)$$

In particular, for any integer a we have

$$(a + 1)^p \equiv a^p + 1 \pmod{p}.$$

It is now immediate to prove by induction on $a \geq 0$ that $a^p \equiv a \pmod{p}$ for all primes p . Similarly (or using that $(-a)^p \equiv -a^p \pmod{p}$) we obtain the result when $a \leq 0$.

Note that Fermat's little theorem and the validity of congruence (1) for all integers x, y are equivalent. Indeed, it is clear that Fermat's little theorem yields congruence (1), since both sides are congruent to $x + y$ modulo p . Conversely, if congruence (1) holds for all integers x, y then a simple inductive argument shows that

$$(x_1 + \dots + x_n)^p \equiv x_1^p + \dots + x_n^p \pmod{p}$$

for all integers x_1, \dots, x_n . In particular, if a is a positive integer then

$$a^p = (\underbrace{1 + \dots + 1}_a)^p \equiv \underbrace{1 + \dots + 1}_a = a \pmod{p}$$

and Fermat's little theorem follows (the case $a < 0$ follows from the case $a \geq 0$ using that $(-a)^p \equiv -a^p \pmod{p}$).

A very important observation concerning Fermat's little theorem is that its converse does **not** hold, in other words there are composite numbers n such that $a^n \equiv a \pmod{n}$ for all integers a . Such numbers are called *Carmichael*

numbers, and the first few are given by $n = 561, 1105, 1729, 2465$. It is known (this is a deep theorem of Alford, Granville and Pomerance) that there are infinitely many Carmichael numbers. The next example explains why the previous numbers are Carmichael numbers.

Example 5.3. Let n be a composite squarefree integer such that $p - 1 \mid n - 1$ for any prime p dividing n . Prove that n is a Carmichael number.

Proof. We need to prove that $a^n \equiv a \pmod{n}$ for any integer a . Since n is squarefree, it suffices to prove that $a^n \equiv a \pmod{p}$ for any prime p dividing n . If $p \mid a$, we are done, otherwise by Fermat's little theorem $a^{p-1} \equiv 1 \pmod{p}$ and since $p - 1 \mid n - 1$ we obtain $a^{n-1} \equiv 1 \pmod{p}$ and then $a^n \equiv a \pmod{p}$, as desired. \square

For instance, $561 = 3 \cdot 11 \cdot 17$ satisfies the conditions imposed in the previous example, since 560 is a multiple of 2, 10 and 16. Thus 561 is a Carmichael number. The argument is similar for $1105 = 5 \cdot 13 \cdot 17$, $1729 = 7 \cdot 13 \cdot 19$, $2465 = 5 \cdot 17 \cdot 29$. We will see later on that the converse holds in the previous example, i.e. any Carmichael number n is squarefree (this is fairly easy to see, since by assumption $n \mid p^n - p$ for any prime $p \mid n$, hence p^2 cannot divide n) and for any prime $p \mid n$ we have $p - 1 \mid n - 1$ (this is difficult to establish using only the tools we have so far).

Example 5.4. Prove that there are infinitely many composite integers n for which $n \mid a^{n-1} - a$ for any integer a .

Proof. We claim that $n = 2p$ with p an odd prime is a solution of the problem. Since $a^{n-1} - a$ is clearly even, it suffices to prove that $p \mid a^{2p-1} - a$ for all a and all odd primes p . This follows from

$$a^{2p-1} - a = a(a^{2p-2} - 1) = (a^p - a)(a^{p-1} + 1)$$

and Fermat's little theorem. \square

Numbers n for which $2^n \equiv 2 \pmod{n}$ are also historically very important. One can show that the first composite number n with this property is $341 = 11 \cdot 31$.

Definition 5.5. A composite integer n such that $2^n \equiv 2 \pmod{n}$ is called a pseudo-prime. More generally, if $a > 1$ is an integer, a composite integer n such that $a^n \equiv a \pmod{n}$ is called a pseudo-prime in base a .

Thus Carmichael numbers are precisely those numbers which are pseudo-primes in any base. The first pseudo-primes are 341, 561, 645, 1105, 1387, 1729, 1905, 2047,.... Combined with the fact that 561 (or 341) is a pseudo-prime, the next example proves the existence of infinitely many pseudo-primes.

Example 5.6. Prove that if n is odd and pseudo-prime, then so is $2^n - 1$.

Proof. Since n is composite, so is $2^n - 1$ (for if d is a proper divisor of n , then $2^d - 1$ is a proper divisor of $2^n - 1$). We need to prove that $2^n - 1 \mid 2^{2^n - 2} - 1$, or equivalently $n \mid 2^n - 2$. But this is clear, since n is a pseudo-prime. \square

The next example gives different proofs of the existence of infinitely many pseudo-primes using explicit constructions.

Example 5.7. a) (Erdős, 1950) Prove that if $p > 3$ is a prime then $n = \frac{4^p - 1}{3}$ is a pseudo-prime.

b) (Rotkiewicz, 1964) Prove that if $p > 5$ is a prime then $n = \frac{4^p + 1}{5}$ is a pseudo-prime.

Proof. a) Note that $n = \frac{2^p - 1}{3} \cdot (2^p + 1)$ is composite. Next, since $n \mid 4^p - 1$, in order to prove that $n \mid 2^n - 2$ it suffices to prove that $4^p - 1 \mid 2^{n-1} - 1$, or equivalently $2p \mid n - 1$. This is further equivalent to $6p \mid 4^p - 4$. Clearly 2 and 3 divide $4^p - 4$ and by Fermat's little theorem $p \mid 4^p - 4$. Since 2, 3, p are pairwise relatively prime, the result follows.

b) Write $p = 2k + 1$, then

$$n = \frac{2^{4k+2} + 1}{5} = \frac{4 \cdot (2^k)^4 + 1}{5} = \frac{(2^{2k+1} - 2^{k+1} + 1)(2^{2k+1} + 2^{k+1} + 1)}{5}$$

and $2^{2k+1} - 2^{k+1} + 1 > 5$ when $p > 5$, hence n is composite. Next, it suffices to prove that $4^p + 1 \mid 2^{n-1} - 1$ and since $4^p + 1 \mid 2^{4p} - 1$ we are further reduced to $4p \mid n - 1$ and then $20p \mid 4^p - 4$. This follows from Fermat's little theorem and the fact that 4, 5, p are pairwise relatively prime and each divides $4^p - 4$. \square

The reader has already noted that all pseudo-primes presented in the above discussion are odd. What about even ones? These are much harder to find: only in 1950 did D.H. Lehmer find the smallest even pseudo-prime, the number $n = 161038 = 2 \cdot 73 \cdot 1103$. To see that n is a pseudo-prime, one uses again Fermat's little theorem and the fact that $n - 1 = 3^2 \cdot 29 \cdot 617$ with $2^9 - 1 = 7 \cdot 73$ and $2^{29} - 1 = 233 \cdot 1103 \cdot 2089$. Beeger proved in 1951 that there are infinitely many even pseudo-primes.

5.1.2 Some concrete examples

We continue with many illustrations of Fermat's little theorem, destined to better grasp the power of this result. We start with a series of interesting congruences that can be derived rather easily using Fermat's little theorem. The trick of considering the smallest prime factor of n when dealing with divisibilities of the form $n \mid a^n - b^n$ is a standard tool which turns out to be very effective in practice. The next two examples illustrate this plainly.

Example 5.8. a) Prove that if $n > 1$, then n does not divide $2^n - 1$.

b) Find all odd positive integers n for which $n \mid 3^n + 1$.

Proof. a) Suppose that $n \mid 2^n - 1$ and let p be the smallest prime divisor of n . Then $p \mid n \mid 2^n - 1$ and by Fermat's little theorem $p \mid 2^{p-1} - 1$. Hence $p \mid \gcd(2^n - 1, 2^{p-1} - 1) = 2^{\gcd(n, p-1)} - 1$. Since p is the smallest prime divisor of n , we have $\gcd(p - 1, n) = 1$, hence $p \mid 1$, a contradiction.

b) The answer is $n = 1$. Suppose that $n > 1$ is a solution and let p be the smallest prime divisor of n . Then $p \mid 3^n + 1 \mid 3^{2n} - 1$ and $p \mid 3^{p-1} - 1$. Thus $p \mid \gcd(3^{2n} - 1, 3^{p-1} - 1) = 3^{\gcd(2n, p-1)} - 1$. Since n is odd, so is p , and since p is the smallest prime divisor of n we have $\gcd(2n, p - 1) = 2$. Thus $p \mid 3^2 - 1 = 8$, a contradiction. \square

Example 5.9. (China TST 2006) Find all positive integers n and all integers a such that $n \mid (a + 1)^n - a^n$.

Proof. Clearly $(n, a) = (1, a)$ is a solution for any integer a . Assume now that $n > 1$ and consider the smallest prime divisor p of n . Then $p \mid (a + 1)^n - a^n$. Note that p cannot divide a or $a + 1$, as otherwise p would divide both a and

$a + 1$. Thus by Fermat's little theorem $p \mid (a + 1)^{p-1} - a^{p-1}$. We deduce that $p \mid (a + 1)^{\gcd(n, p-1)} - a^{\gcd(n, p-1)}$ and since $\gcd(n, p-1) = 1$ it follows that $p \mid 1$, a contradiction. Thus we have already found all solutions. \square

For the next example, we recall that $v_p(n)$ denotes the exponent of p in the prime factorization of n .

Example 5.10. a) Let n be a positive integer and let p be a prime factor of $2^n + 1$. Prove that $v_2(p-1) > v_2(n)$.

b) Find all prime numbers p, q such that $pq \mid 2^p + 2^q$.

Proof. a) We have $p \mid 2^{2^n} - 1$ and $p \mid 2^{p-1} - 1$, thus $p \mid \gcd(2^{2^n} - 1, 2^{p-1} - 1) = 2^{\gcd(2n, p-1)} - 1$. Suppose that $v_2(p-1) \leq v_2(n)$, then $\gcd(2n, p-1) \mid n$ and we conclude that $p \mid 2^n - 1$. Since $p \mid 2^n + 1$, it follows that $p \mid 2$, a contradiction. Hence $v_2(p-1) > v_2(n)$.

b) If $p = 2$ then $2q \mid 4 + 2^q$. Since $4 + 2^q \equiv 6 \pmod{q}$ by Fermat's little theorem, we deduce that $q \mid 6$ and so $q = 2$ or $q = 3$, both of which are solutions of the problem. By symmetry if $q = 2$ then $p = 2$ or $p = 3$. Assume now that $p, q > 2$ and without loss of generality assume that $p > q$. Then by assumption $pq \mid 2^{p-q} + 1$. It follows from part a) that $v_2(p-1) > v_2(p-q)$ and $v_2(q-1) > v_2(p-q)$. This is impossible, since

$$v_2(p-q) = v_2((p-1) - (q-1)) \geq \min(v_2(p-1), v_2(q-1)).$$

Thus the only solutions of the problem are $(2, 2), (2, 3), (3, 2)$. \square

Example 5.11. Let $(f_n)_{n \geq 1}$ be the Fibonacci sequence, with $f_1 = f_2 = 1$ and $f_{n+1} = f_n + f_{n-1}$ for $n \geq 2$. Prove that for any prime $p > 2$ we have

$$f_p \equiv 5^{\frac{p-1}{2}} \pmod{p}.$$

Proof. We use the classical formula

$$f_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right),$$

which can be established by a straightforward induction. Expanding the expression on the right-hand side using the binomial formula yields

$$f_p = \frac{1}{2^p \sqrt{5}} \sum_{k=0}^p \binom{p}{k} 5^{\frac{k}{2}} (1 - (-1)^k) = \frac{1}{2^{p-1}} \sum_{k=0}^{\frac{p-1}{2}} \binom{p}{2k+1} 5^k.$$

Since p divides $\binom{p}{2k+1}$ for $0 \leq k \leq \frac{p-3}{2}$ we deduce that

$$2^{p-1} f_p \equiv 5^{\frac{p-1}{2}}$$

and since $2^{p-1} \equiv 1 \pmod{p}$ by Fermat's little theorem, the result follows. \square

Example 5.12. Prove that for all odd primes p we have

$$\sum_{k=1}^{p-1} k^{2p-1} \equiv \frac{p(p+1)}{2} \pmod{p^2}.$$

Proof. By Fermat's little theorem we have $k(k^{p-1} - 1)^2 \equiv 0 \pmod{p^2}$. Expanding this and summing we find

$$\sum_{k=1}^{p-1} k^{2p-1} \equiv 2 \sum_{k=1}^{p-1} k^p - \sum_{k=1}^{p-1} k \pmod{p^2}.$$

On the other hand,

$$2 \sum_{k=1}^{p-1} k^p = \sum_{k=1}^{p-1} (k^p + (p-k)^p) \equiv 0 \pmod{p^2}$$

since $k^p + (p-k)^p \equiv 0 \pmod{p^2}$ for $1 \leq k \leq p-1$ (as follows directly from the binomial formula). We conclude that

$$\sum_{k=1}^{p-1} k^{2p-1} \equiv - \sum_{k=1}^{p-1} k = -\frac{p(p-1)}{2} \equiv \frac{p(p+1)}{2} \pmod{p^2}. \quad \square$$

Fermat's little theorem can be very efficient in establishing that certain numbers are composite or in proving that certain sequences contain infinitely many composite numbers, as shown in the following examples.

Example 5.13. Let $a_1, \dots, a_n, b_1, \dots, b_k$ be integers such that $a_1, \dots, a_n > 1$. Prove that there are infinitely many positive integers d such that $a_1^d + a_2^d + \dots + a_n^d + b_i$ is composite for all $1 \leq i \leq k$.

Proof. Since $a_1, \dots, a_n > 1$, there is a positive integer d such that

$$S_i := a_1^d + \dots + a_n^d + b_i > 1$$

for $1 \leq i \leq k$. Let p_i be a prime divisor of S_i and let $d_j = d + j(p_1 - 1) \dots (p_k - 1)$. By Fermat's little theorem

$$a_1^{d_j} + \dots + a_n^{d_j} + b_i \equiv a_1^d + \dots + a_n^d + b_i \equiv 0 \pmod{p_i}$$

for any $j > 1$ and since clearly $a_1^{d_j} + \dots + a_n^{d_j} + b_i > S_i \geq p_i$ for $j \geq 1$ and $1 \leq i \leq k$, it follows that $a_1^{d_j} + \dots + a_n^{d_j} + b_i$ is composite for $1 \leq i \leq k$ and $j \geq 1$. \square

Example 5.14. (China TST 2002) Are there distinct positive integers k_1, \dots, k_{2002} such that for all integers $n > 2001$ at least one of the numbers $k_1 \cdot 2^n + 1, \dots, k_{2002} \cdot 2^n + 1$ is prime?

Proof. The answer is negative: choose a prime divisor p_i of $2k_i + 1$ for $1 \leq i \leq 2002$, and let $n = N(p_1 - 1) \dots (p_{2002} - 1) + 1$ for some large $N > 2001$. Then $n > 2001$ and by Fermat's little theorem $k_i \cdot 2^n + 1 \equiv 2k_i + 1 \equiv 0 \pmod{p_i}$ for $1 \leq i \leq 2002$. Moreover, it is clear that $k_i \cdot 2^n + 1 > p_i$, hence $k_i \cdot 2^n + 1$ is composite for $1 \leq i \leq 2002$. \square

Example 5.15. Let $k > 1$ be integer and define $a_n = 2^{2^n} + k$. Prove that there are infinitely many composite numbers in the sequence a_1, a_2, \dots .

Proof. The solution is short, but fairly tricky. We may assume that k is odd, since otherwise all terms of the sequence are even. Let $r = v_2(k - 1)$ (where $v_2(x)$ is the exponent of 2 in the prime factorization of x) and assume that a_n is prime for all large enough n , say $n > N$. In particular there is $n > \max(r, N)$ such that a_n is a prime number, say $a_n = p$. Since $n > r$ we have $v_2(p - 1) = v_2(2^{2^n} + k - 1) = r$. Write $p - 1 = 2^r \cdot s$ for some odd number s and choose a positive integer j such that $2^j \equiv 1 \pmod{s}$ (to see

that this is possible, follow the proof of corollary 4.15 or use Euler's theorem in chapter 6). Then $2^{j+n} \equiv 2^n \pmod{p-1}$ and so by Fermat's little theorem $2^{2^{j+n}} + k \equiv a_n \equiv 0 \pmod{p}$. Thus a_{j+n} is divisible by p and since clearly $a_{j+n} > a_n = p$ we deduce that a_{j+n} is composite, a contradiction. \square

The next examples are concerned with various divisibility properties that can be deduced from Fermat's little theorem, with a special emphasis on polynomials.

Example 5.16. (Poland) Find all polynomials f with integer coefficients such that $f(n) | 2^n - 1$ for all positive integers n .

Proof. Clearly the constant polynomials 1 and -1 are solutions of the problem. Conversely, let f be a solution of the problem and suppose that $f(n)$ is not ± 1 for some n . Then $f(n)$ must have a prime factor p . Then p divides $f(n+p) | 2^{n+p} - 1$ and p divides $f(n) | 2^n - 1$. We conclude that $p | 2^p - 1$, contradicting Fermat's little theorem. Thus $f(n) = \pm 1$ for all n , which immediately implies that f is a constant polynomial, equal to 1 or -1 . \square

Example 5.17. (ELMO 2016) Let f be a polynomial with integer coefficients such that $n | f(2^n)$ for all $n \geq 1$. Prove that $f = 0$.

Proof. If p, q are distinct odd primes, then by assumption $pq | f(2^{pq})$, thus $f(2^{pq}) \equiv 0 \pmod{p}$. On the other hand, Fermat's little theorem yields $2^{pq} \equiv 2^q \pmod{p}$, thus $f(2^{pq}) \equiv f(2^q) \pmod{p}$. We conclude that $p | f(2^q)$ for any distinct odd primes p, q . Fixing $q > 2$ and letting p vary, it follows that $f(2^q) = 0$. We conclude that f has infinitely many zeros and so $f = 0$. \square

Example 5.18. Let $p \geq 5$ be a prime and let a, b be integers such that p divides $a^2 + ab + b^2$. Prove that

$$(a+b)^p \equiv a^p + b^p \pmod{p^2}.$$

Proof. If $p | a$, then $p | b$ and the result is clear. So assume that p does not divide ab . Let x be an integer such that $bx \equiv a \pmod{p^2}$, then $p | x^2 + x + 1$ and so $p | x^3 - 1$. Using the binomial formula

$$x^{3p} - 1 = (x^3 - 1 + 1)^p - 1 = (x^3 - 1)^p + \dots + p(x^3 - 1)$$

we deduce that $p^2 \mid x^{3p} - 1$ and so $p^2 \mid (x^p - 1)(x^{2p} + x^p + 1)$. On the other hand, p does not divide $x^p - 1$, since otherwise, by Fermat's little theorem, p would divide $x - 1$. Since it also divides $x^2 + x + 1$, we would have $p \mid 3$, a contradiction. Thus $p^2 \mid x^{2p} + x^p + 1$. On the other hand, since $x + 1 \equiv -x^2 \pmod{p}$, we have $(x + 1)^p \equiv -x^{2p} \pmod{p^2}$. Combining these results yields

$$(x + 1)^p \equiv x^p + 1 \pmod{p^2}.$$

The result follows by multiplying this congruence by b^p and using that $bx \equiv a \pmod{p^2}$. \square

Remark 5.19. A stronger result holds: the congruence holds modulo p^3 , but the proof is different. One proves that $p(X^2 + X + 1)^2$ divides the polynomial $(X + 1)^p - X^p - 1$ in $\mathbf{Z}[X]$.

The last series of examples concerns exponential sequences and congruences.

Example 5.20. a) Prove that for any prime $p > 2$ there are infinitely many positive integers n such that $n \cdot 2^n + 1 \equiv 0 \pmod{p}$.

b) (IMO 2005) Which positive integers are relatively prime to all numbers of the form $2^n + 3^n + 6^n - 1$, with $n \geq 1$?

Proof. a) We choose $n = k(p - 1) + r$ with $k \geq 1$ and $r \geq 0$. Then

$$n \cdot 2^n + 1 \equiv (r - k)2^r + 1 \pmod{p}$$

by Fermat's little theorem. It is thus enough to ensure that $p \mid (r - k)2^r + 1$. Simply choose $r = 0$ and $k \equiv 1 \pmod{p}$.

b) We will prove that 1 is the unique solution of the problem, by showing that for any prime p there is $n \geq 1$ such that $p \mid a_n$. Note that 2 and 3 divide $a_2 = 48$, hence we may assume that $p > 3$. Then using Fermat's little theorem we obtain

$$6a_{p-2} = 3 \cdot 2^{p-1} + 2 \cdot 3^{p-1} + 6^{p-1} - 6 \equiv 3 + 2 + 1 - 6 \equiv 0 \pmod{p}.$$

Since $\gcd(6, p) = 1$ it follows that $a_{p-2} \equiv 0 \pmod{p}$, thus $p \mid a_{p-2}$ and the problem is solved. \square

Example 5.21. (IMO Shortlist 2005) Let a, b be positive integers such that $a^n + n$ divides $b^n + n$ for all positive integers n . Prove that $a = b$.

Proof. Pick a large prime $p > \max(a, b)$ and let us look for n such that $p \mid a^n + n$. Choosing $n = (p-1)k + r$ for suitable k, r , we have by Fermat's little theorem $a^n + n \equiv a^r - k + r \pmod{p}$, so it suffices to take any positive integer r and $k = a^r + r$. With this choice we have $p \mid a^n + n \mid b^n + n$ and again by Fermat's little theorem

$$b^n + n \equiv b^r + r - k = b^r - a^r \pmod{p}.$$

We deduce that $p \mid b^r - a^r$ for any prime $p > b$ and any positive integer r . Choosing $r = 1$ we obtain $a = b$. \square

Example 5.22. (Komal) Let $p_1 = 2$ and p_{n+1} be the smallest prime divisor of the number $np_1^{1!}p_2^{2!}\dots p_n^{n!} + 1$. Prove that every prime number appears in the sequence p_1, p_2, \dots

Proof. To simplify notations, write $x_n = np_1^{1!}\dots p_n^{n!}$. Since $p_{n+1} \mid x_n + 1$ and $p_1 \dots p_n \mid x_n$, it is clear that p_{n+1} is different from any of p_1, \dots, p_n and so the terms of the sequence are pairwise distinct. It remains to prove that any prime appears in the sequence. Suppose that this is not the case and let p be the smallest prime number which does not appear in the sequence. Take $n > p$ large enough so that all primes less than p are among p_1, \dots, p_n . Then for any $k \geq 1$ we have

$$x_{n+k} \equiv (n+k)p_1^{1!}\dots p_{p-2}^{(p-2)!} \pmod{p}$$

since $p-1 \mid j!$ for $j \geq p-1$ and by Fermat's little theorem $p_j^{j!} \equiv 1 \pmod{p}$ for such j (note that by assumption $p \neq p_j$ so $\gcd(p, p_j) = 1$). Since p is relatively prime to $p_1^{1!}\dots p_{p-2}^{(p-2)!}$, we can choose k such that $(n+k)p_1^{1!}\dots p_{p-2}^{(p-2)!} + 1 \equiv 0 \pmod{p}$, thus $p \mid x_{n+k} + 1$. Any prime less than p already divides $x_{n+k} + 1$, so p is the smallest prime factor of $x_{n+k} + 1$. It follows that $p = p_{n+k+1}$, a contradiction. \square

Example 5.23. (Romanian Masters in Mathematics 2012) Prove that there are infinitely many positive integers n such that n divides $2^{2^n+1} + 1$ but it does not divide $2^n + 1$.

Proof. For each $k \geq 1$ let $a_k = 2^{3^k} + 1$. Observe that

$$a_{k+1} = (a_k - 1)^3 + 1 = a_k(a_k^2 - 3a_k + 3),$$

which immediately yields by induction that $3^{k+1} \mid a_k$ and so the number

$$b_k = \frac{a_k^2 - 3a_k + 3}{3} = a_k \cdot \frac{a_k}{3} - a_k + 1$$

is an integer greater than 1 (since $a_k > 3$) and relatively prime to a_k . Let p_k be a prime divisor of b_k . Note that $p_k \mid a_{k+1}$ but p_k does not divide a_k .

Define $n_k = 3^k \cdot p_k$. Then by Fermat's little theorem

$$2^{n_k} + 1 = (2^{3^k})^{p_k} + 1 \equiv 2^{3^k} + 1 = a_k \pmod{p_k},$$

thus p_k does not divide $2^{n_k} + 1$, in particular n_k does not divide $2^{n_k} + 1$. Next, we claim that $n_k \mid 2^{2^{n_k}+1} + 1$. Since $n_k \mid a_{k+1}$, it suffices to prove that $a_{k+1} \mid 2^{2^{n_k}+1} + 1$, or equivalently that $3^{k+1} \mid 2^{n_k} + 1$. But $2^{n_k} + 1$ is a multiple of $2^{3^k} + 1 = a_k$, which in turn is a multiple of 3^{k+1} , so we are done. \square

Remark 5.24. We leave it as an easy exercise for the reader to prove that if n has the given property then so does $2^n + 1$. This gives an alternative solution as soon as we are able to exhibit at least one such n . It is not difficult to check that $n = 57$ is such a number.

Example 5.25. (Russia 2013) Find all positive integers k for which there exist positive integers a and $n > 1$ such that $a^n + 1$ is the product of the first k odd primes.

Proof. We will prove that no such k exists. Assume by contradiction that $a^n + 1 = p_1 p_2 \dots p_k$, where $p_1 = 3, p_2 = 5, \dots$ is the increasing sequence of odd primes. Clearly $k > 1$. Note that since 3 divides $a^n + 1$, n must be odd. Next, we will prove that $a \leq p_k$. Suppose that $a > p_k$, then since $a^n + 1 < p_k^k$, we must have $n < k$ and in particular $n < p_k$. Let p be a prime factor of n , then $p \in \{p_1, \dots, p_k\}$. Moreover, p divides $a^n + 1$, hence if we let $b = a^{n/p}$, we have $p \mid b^p + 1$. Fermat's little theorem yields $p \mid b + 1$. But then $p^2 \mid b^p + 1 = a^n + 1$ since $b^{p-1} - b^{p-2} + \dots + 1 \equiv 0 \pmod{p}$. This contradicts the fact that $a^n + 1$ is square free and finishes the proof of the claim that $a \leq p_k$.

Next, assume that $a > 2$ and let p be a prime factor of $a - 1$. Then $a^n + 1 \equiv 2 \pmod{p}$, hence $p \notin \{p_1, \dots, p_k\}$ and so $a > p_k$, a contradiction. Thus $a = 2$. Since $5 \mid 2^n + 1$, n must be even, contradiction again! \square

Example 5.26. (China TST 2008) Let n be an integer greater than 1 such that n divides $2^{\varphi(n)} + 3^{\varphi(n)} + \dots + n^{\varphi(n)}$. If p_1, \dots, p_k are all the prime divisors of n (without multiplicities), prove that $\frac{1}{p_1} + \frac{1}{p_2} + \dots + \frac{1}{p_k} + \frac{1}{p_1 p_2 \dots p_k}$ is an integer.

Proof. Fix $i \in \{1, 2, \dots, k\}$. By assumption p_i divides $2^{\varphi(n)} + 3^{\varphi(n)} + \dots + n^{\varphi(n)}$. If $a \in \{2, 3, \dots, n\}$, then either a is a multiple of p_i , and then $p_i \mid a^{\varphi(n)}$, or not, and then $a^{\varphi(n)} \equiv 1 \pmod{p_i}$ (by Fermat's little theorem and the fact that $p_i - 1$ divides $\varphi(n)$). Hence $2^{\varphi(n)} + 3^{\varphi(n)} + \dots + n^{\varphi(n)}$ is congruent modulo p_i to the number of $a \in \{2, 3, \dots, n\}$ which are not multiples of p_i . This number is $n - 1 - \frac{n}{p_i}$ and since $p_i \mid n$, it follows that $p_i \mid \frac{n}{p_i} + 1$. In particular p_i^2 does not divide n , and so $n = p_1 p_2 \dots p_k$. Moreover, p_i divides $\prod_{j \neq i} p_j + 1$ for all i . It follows that $p_2 \dots p_k + p_1 p_3 \dots p_k + \dots + p_1 \dots p_{k-1} + 1$ is a multiple of p_1, \dots, p_k , thus also a multiple of $p_1 p_2 \dots p_k$. But this is precisely saying that $\frac{1}{p_1} + \frac{1}{p_2} + \dots + \frac{1}{p_k} + \frac{1}{p_1 p_2 \dots p_k}$ is an integer. \square

5.1.3 Application to primes of the form $4k + 3$ and $3k + 2$

After this series of examples, we come back for a while to more theoretical issues. The first result shows that n th powers modulo p are solutions to the congruence $x^{\frac{p-1}{n}} \equiv 1 \pmod{p}$ whenever $n \mid p - 1$. We will see later on that all solutions of this congruence are n th powers modulo p .

Proposition 5.27. Let p be a prime and let n be a positive integer dividing $p - 1$. If a is an integer such that the congruence $x^n \equiv a \pmod{p}$ has solutions (in other words a is an n th power modulo p), then $p \mid a$ or $a^{\frac{p-1}{n}} \equiv 1 \pmod{p}$.

Proof. This is an immediate application of Fermat's little theorem: if p does not divide a , then

$$a^{\frac{p-1}{n}} \equiv (x^n)^{\frac{p-1}{n}} = x^{p-1} \equiv 1 \pmod{p}$$

and we are done. \square

The previous proposition easily yields the following result, which is very useful in practice. We will see later on that it characterizes primes of the form $4k + 3$.

Corollary 5.28. *Let p be a prime of the form $4k + 3$. If $p \mid a^2 + b^2$ for some integers a, b then $p \mid a$ and $p \mid b$.*

Proof. If $p \mid a$ then clearly $p \mid b^2$ and so $p \mid b$. Assume now that p does not divide a and let c be an integer such that $ac \equiv 1 \pmod{p}$. Since $p \mid (ac)^2 + (bc)^2$, we obtain $(bc)^2 \equiv -1 \pmod{p}$ and by the previous proposition $(-1)^{\frac{p-1}{2}} \equiv 1 \pmod{p}$. Since $p \equiv 3 \pmod{4}$, the last congruence reads $-1 \equiv 1 \pmod{p}$, plainly absurd. \square

The following theorem is also very useful in practice.

Theorem 5.29. *Let p be a prime and let n be a positive integer relatively prime to $p - 1$. Then the remainders of $1^n, 2^n, \dots, (p - 1)^n$ when divided by p are a permutation of $1, 2, \dots, p - 1$.*

Proof. Clearly none of these numbers is a multiple of p . It suffices therefore to prove that the numbers are pairwise incongruent modulo p . Suppose that $p \mid a^n - b^n$ for some $a, b \in \{1, 2, \dots, p - 1\}$ and note that we may assume that $\gcd(a, b) = 1$ (since p does not divide $\gcd(a, b)$). Then using Fermat's little theorem we also have $p \mid a^{p-1} - b^{p-1}$ and so (using proposition 3.35)

$$p \mid \gcd(a^n - b^n, a^{p-1} - b^{p-1}) = a^{\gcd(n, p-1)} - b^{\gcd(n, p-1)} = a - b,$$

the last equality being a consequence of our hypothesis. Since $p \mid a - b$ and $a, b \in \{1, 2, \dots, p - 1\}$, we must have $a = b$ and we are done. \square

Corollary 5.30. *Let p be a prime of the form $3k + 2$. Then*

a) *The remainders of the numbers $1^3, 2^3, \dots, (p - 1)^3$ when divided by p are a permutation of $1, 2, \dots, p - 1$.*

b) *If $p \mid a^2 + ab + b^2$ for some integers a, b , then $p \mid a$ and $p \mid b$.*

c) *If $p \neq 2$ then there is no integer x such that $x^2 \equiv -3 \pmod{p}$.*

Proof. a) This follows directly from theorem 5.29 for $n = 3$.

b) If $p \mid a$ then $p \mid b$, so assume that p does not divide ab . Then $p \mid (a - b)(a^2 + ab + b^2) = a^3 - b^3$ and by part a) we deduce that $p \mid a - b$. But since $p \mid a^2 + ab + b^2$, it follows that $p \mid 3a^2$, a contradiction. The result follows.

c) Suppose that x is such an integer. Since $p \neq 2$, there is an integer y such that $2y + 1 \equiv x \pmod{p}$, then $4y^2 + 4y + 4 \equiv 0 \pmod{p}$ and so $y^2 + y + 1 \equiv 0 \pmod{p}$. But this contradicts part b). \square

Example 5.31. Prove that there are infinitely many primes of the form $4k + 1$ and infinitely many primes of the form $6k + 1$.

Proof. By Schur's theorem 4.67 there are infinitely many primes p dividing a number of the form $n^2 + 1$ with $n \geq 1$. Corollary 5.28 shows that any such p is either equal to 2 or of the form $4k + 1$. We deduce the first part of the problem. For the second part consider similarly prime divisors of numbers of the form $n^2 + n + 1$ with $n \geq 1$. Corollary 5.30 shows that such primes are of the form $3k + 1$ (thus of the form $6l + 1$) or equal to 3. The result follows. \square

Example 5.32. Find all integers a and b such that $a^2 - 1 \mid b^2 + 1$.

Proof. Clearly $(a, b) = (0, n)$ works for all integers n , and we will prove that these are all solutions. So, suppose that (a, b) is a solution with $a \neq 0$. Then clearly $a \neq \pm 1$, hence $a^2 - 1 > 1$. If a is odd, then 8 divides $a^2 - 1$, hence $8 \mid b^2 + 1$, which is impossible. Hence a is even, thus $a^2 - 1 \equiv 3 \pmod{4}$. Since $a^2 - 1 > 1$, it follows that $a^2 - 1$ has a prime factor p of the form $4k + 3$. But p cannot divide $b^2 + 1$, a contradiction. This finishes the proof. \square

Example 5.33. Prove that if a is an integer, then $2a^2 - 1$ has no divisors of the form $b^2 + 2$ with $b \in \mathbb{Z}$.

Proof. Suppose that $b^2 + 2 \mid 2a^2 - 1$ for some integers a, b . Then clearly b is odd, thus $b^2 + 2 \equiv 3 \pmod{4}$. It follows that $b^2 + 2$ has a prime factor p of the form $4k + 3$. Then $p \mid b^2 + 2$ and $p \mid 2a^2 - 1$, thus

$$p \mid b^2 + 2 + 2(2a^2 - 1) = b^2 + (2a)^2.$$

It follows that $p \mid b$ and $p \mid 2a$, which is clearly impossible. \square

Example 5.34. (Iran 2004) Find all primes p, q, r such that $p^3 = p^2 + q^2 + r^2$.

Proof. If p, q, r are not multiples of 3 then $p^2 + q^2 + r^2 \equiv 1 + 1 + 1 \equiv 0 \pmod{3}$ and so $3 \mid p^3$, a contradiction. Hence one of p, q, r is 3. If $p = 3$ then $q^2 + r^2 = 18$, which easily yields $q = r = 3$. Assume that $p > 3$ and without loss of generality that $r = 3$, hence $p^3 = p^2 + q^2 + 9$, that is $p^2(p-1) = q^2 + 9$. If $p \equiv 1 \pmod{4}$, we deduce that $4 \mid q^2 + 9$, thus $4 \mid q^2 + 1$, which is impossible. Thus $p \equiv 3 \pmod{4}$. But since $p \mid q^2 + 3^2$, we obtain $p \mid q$ and $p \mid 3$, thus $p = q = 3$ and then $r = 3$. \square

Example 5.35. (Brazil 1996) Let $P(x) = x^3 + 14x^2 - 2x + 1$ and let $P^{[n]}$ be the composition of P with itself n times (so $P^{[3]}(x) = P(P(P(x)))$). Prove that there is a positive integer n such that $P^{[n]}(x) \equiv x \pmod{101}$ for all integers x .

Proof. Let $p = 101$. Define the function $f : \{0, 1, \dots, p-1\} \rightarrow \{0, 1, \dots, p-1\}$ by setting $f(i)$ the remainder of $P(i)$ when divided by p . We need to prove that there is $n \geq 1$ such that $f^{[n]}$ is the identity map. This is equivalent to saying that f is bijective: indeed, it is clear that the existence of n forces f being bijective, so suppose that f is bijective. Since there are finitely many maps $g : \{0, 1, \dots, p-1\} \rightarrow \{0, 1, \dots, p-1\}$, the sequence of iterates $f, f^{[2]}, f^{[3]}, \dots$ cannot consist of pairwise different functions. So there are $0 \leq i < j$ such that $f^{[i]} = f^{[j]}$ and we can choose $n = j - i$.

Now, in order to prove that f is bijective, it suffices to prove its injectivity (since the source and target of f have the same number of elements). But if $f(i) = f(j)$ then

$$p \mid P(i) - P(j) = (i - j)(i^2 + ij + j^2 + 14(i + j) - 2).$$

Assume that $i \neq j$ then $p \mid i^2 + ij + j^2 + 14(i + j) - 2$. Choose α such that $3\alpha \equiv 14 \pmod{p}$ and note that

$$(i + \alpha)^2 + (i + \alpha)(j + \alpha) + (j + \alpha)^2 \equiv i^2 + ij + j^2 + 14(i + j) + 3\alpha^2 \equiv 3\alpha^2 + 2.$$

But $9\alpha^2 \equiv 14^2 = 196 \equiv -6 \pmod{p}$ and so $p \mid 3\alpha^2$. It follows that

$$(i + \alpha)^2 + (i + \alpha)(j + \alpha) + (j + \alpha)^2 \equiv 0 \pmod{p}$$

and since $p \equiv 2 \pmod{3}$ we deduce that $p \mid i + \alpha$ and $p \mid j + \alpha$. Thus $p \mid i - j$ and then $i = j$, a contradiction. \square

Remark 5.36. One can replace $p = 101$ with any prime congruent to 2 modulo 3 and P with any polynomial of the form $P(x) = x^3 + ax^2 + bx + c$ with $a^2 \equiv 3b \pmod{p}$.

Example 5.37. (IMO Shortlist 2012) Find all triples (x, y, z) of positive integers such that

$$x^3(y^3 + z^3) = 2012(xyz + 2).$$

Proof. Note that $2012 = 4p$, where $p = 503$ is a prime of the form $3k + 2$. If $p \mid x$, then p^2 divides the left-hand side, while the right-hand side is congruent to $8p$ modulo p^2 , a contradiction. Thus p does not divide x and so $p \mid y^3 + z^3 = y^3 - (-z)^3$. Since $p \equiv 2 \pmod{3}$, it follows that $p \mid y - (-z) = y + z$. Next, $x^3 \mid 4p(xyz + 2)$, thus $x \mid 8p$ and since $\gcd(p, x) = 1$ we obtain $x \mid 8$. If $4 \mid x$, then the left-hand side is a multiple of 16, while the right-hand side is not. Thus $x \in \{1, 2\}$.

Suppose first that $x = 1$, so $y^3 + z^3 = 4p(yz + 2)$. Clearly $2 \mid y + z$, so $2p \mid y + z$. Write the equation as

$$\frac{y+z}{2p} \cdot (y^2 - yz + z^2) = 2(yz + 2).$$

If $\frac{y+z}{2p} = 1$ then $y^2 - 3yz + z^2 = 4$ and so $(y+z)^2 - 5yz = 4$, yielding $p^2 \equiv 1 \pmod{5}$, a contradiction. Thus $\frac{y+z}{2p} \geq 2$ and then $yz + 2 \geq y^2 - yz + z^2$, that is $(y-z)^2 \leq 2$. Since moreover $y \equiv z \pmod{2}$, we deduce that $y = z$ and then $y^3 = 2p(y^2 + 2)$. Since $p \mid y$, taking the last equation modulo p^2 yields a contradiction. Hence the case $x = 1$ is impossible.

Assume now that $x = 2$, then the equation becomes

$$\frac{y+z}{p} \cdot (y^2 - yz + z^2) = yz + 1.$$

Since $p \mid y + z$, we obtain $yz + 1 \geq y^2 - yz + z^2$ and so $(y-z)^2 \leq 1$. If $y = z$ we obtain $\frac{2y}{p} \cdot y^2 = y^2 + 1$ and so $y^2 \mid 1$, giving no solution. Thus, by symmetry, we may assume that $y - z = 1$ and then the equation becomes $y + z = p$, that

is $z = \frac{p-1}{2} = 251$ and $y = 252$. Hence the only solutions are $(2, 251, 252)$ and $(2, 252, 251)$. \square

Example 5.38. (Turkey TST 2013) Find all pairs of positive integers (m, n) such that

$$m^6 = n^{n+1} + n - 1.$$

Proof. If $n = 1$ then $m = 1$, which gives a solution of the problem. One easily checks that $n = 2$ does not yield any solution, so assume that $n > 2$ and that we can find $m > 0$ such that $m^6 = n^{n+1} + n - 1$. Let $k = n + 1 > 2$ and write the equation as

$$m^6 = (k - 1)^k + k - 2.$$

If k is even, then $m^6 > (k - 1)^k$ yields $m^3 \geq (k - 1)^{\frac{k}{2}} + 1$ and then

$$k - 2 \geq 2(k - 1)^{\frac{k}{2}} + 1 > 2(k - 1) + 1,$$

a contradiction. A similar argument (using that m^6 is a third power) shows that 3 does not divide k .

Suppose that $k \equiv 1 \pmod{3}$, then $m^6 \equiv -1 \pmod{3}$, a contradiction. Hence $k \equiv 2 \pmod{3}$ and since k is odd it follows that there is a prime $p > 2$ of the form $3j + 2$ dividing k . Taking the equation mod p yields $m^6 \equiv -3 \pmod{p}$. However this contradicts corollary 5.30(c), and so the equation has no solution except $(m, n) = (1, 1)$. \square

Example 5.39. (Kolmogorov Cup) Let a, b, c be positive integers such that $\frac{a^2+b^2+c^2}{ab+bc+ca}$ is an integer. Prove that this integer is not a multiple of 3.

Proof. Suppose that $a^2 + b^2 + c^2 = 3n(ab + bc + ca)$ for some positive integer n , then

$$(a + b + c)^2 = (3n + 2)(ab + bc + ca).$$

Dividing a, b, c by their greatest common divisor, we may assume that $\gcd(a, b, c) = 1$. Let $3n + 2 = p_1^{\alpha_1} \dots p_k^{\alpha_k}$ be the prime factorization of $3n + 2$ and note that there is i such that $p_i \equiv 2 \pmod{3}$ and α_i is odd, otherwise $p_i^{\alpha_i} \equiv 1 \pmod{3}$ for all i and $3n + 2 \equiv 1 \pmod{3}$, absurd. Fix such i , then $p_i \mid a + b + c$ and since the exponent of p_i in the prime factorization of $(a + b + c)^2$

is even and that in the prime factorization of $3n + 2$ is odd, it follows that $p_i \mid ab + bc + ca$. But then

$$0 \equiv ab + bc + ca \equiv ab + c(a + b) \equiv ab - (a + b)^2 = -(a^2 + ab + b^2) \pmod{p_i}$$

and since $p_i \equiv 2 \pmod{3}$, we deduce that $p_i \mid a$ and $p_i \mid b$, then $p_i \mid c$. This contradicts the relation $\gcd(a, b, c) = 1$ and finishes the proof. \square

5.2 Wilson's theorem

5.2.1 Wilson's theorem as criterion of primality

While Fermat's theorem gives a result that is true for all primes, it does not provide a conclusive test of primality. Wilson's theorem gives an exact criterion for the primality of an integer. The reader is strongly advised to carefully study the proof of the following theorem, since variations on this idea will be encountered several times later on.

Theorem 5.40. (*Wilson's theorem*) a) For all primes p we have

$$(p - 1)! + 1 \equiv 0 \pmod{p}.$$

b) Conversely, if an integer $n > 1$ satisfies $(n - 1)! + 1 \equiv 0 \pmod{n}$, then n is a prime.

Proof. a) For each $i \in \{1, 2, \dots, p - 1\}$ let i^{-1} be the inverse of i modulo p (recall that this is the unique number x between 1 and $p - 1$ which satisfies $ix \equiv 1 \pmod{p}$). We can make a partition of $\{1, 2, \dots, p - 1\}$ into pairs and singletons as follows: pair each i with i^{-1} , if $i \neq i^{-1}$, otherwise put i in a singleton. The product of elements in each pair is 1 modulo p , hence $(p - 1)! = 1 \cdot 2 \cdot \dots \cdot (p - 1)$ is congruent to the product of the numbers in the singletons. However, saying that i lives in a singleton is the same as saying that $i^2 \equiv 1 \pmod{p}$, which is the same as $(i - 1)(i + 1) \equiv 0 \pmod{p}$. Since p is a prime, this is equivalent to $i \equiv \pm 1 \pmod{p}$. Hence there are only two singletons, and the product of their elements is -1 . The result follows.

b) Suppose that n is composite and write $n = ab$ with $a, b > 1$. Then $ab - 1 \geq a$, hence $a \mid (n - 1)!$. By hypothesis $a \mid n \mid (n - 1)! + 1$, hence $a \mid 1$, a contradiction. Hence n is a prime. \square

We illustrate the previous theorem with a few examples.

Example 5.41. (Baltic Way 2014) Is $712! + 1$ a prime number?

Proof. One easily checks that 719 is a prime number, thus Wilson's theorem yields $718! + 1 \equiv 0 \pmod{719}$. Since $718! \equiv 712! \cdot 6! \pmod{719}$ and $6! = 720 \equiv 1 \pmod{719}$, we obtain $719 | 712! + 1$, which shows that $712! + 1$ is composite. \square

Example 5.42. (USAMO 2012) Find all functions $f : \mathbf{N} \rightarrow \mathbf{N}$ such that for all positive integers m, n we have $m - n \mid f(m) - f(n)$ and $f(n!) = f(n)!$.

Proof. The only solutions in positive integers of the equation $n = n!$ are $n = 1, 2$, so the only constant functions which are solutions of the problem are 1, 2. Let f be a nonconstant solution. Since $f(1) = f(1)!$ and $f(2) = f(2)!$, we deduce that $f(1), f(2) \in \{1, 2\}$. If p is an odd prime, then Wilson's theorem combined with the hypothesis yield

$$p \mid (p-2)! - 1 \mid f((p-2)!) - f(1) = f(p-2)! - f(1).$$

Since $f(1) \in \{1, 2\}$ we deduce that p does not divide $f(p-2)!$ and so $f(p-2) \leq p-1$ for all odd primes p . Suppose that $f(p-2) = p-1$ for some $p > 2$, then $p \mid (p-1)! - f(1)$ and by Wilson's theorem again $p \mid f(1) + 1$ thus $p \mid 6$. We deduce that if $p > 3$, then $f(p-2) \leq p-2$. Since moreover $(p-2)! - 1 \leq f(p-2)! - f(1)$, it follows that $f(1) = 1$ and $f(p-2) = p-2$ for all primes $p > 3$. Now, if n is any positive integer then $n - (p-2) \mid f(n) - f(p-2) = f(n) - (p-2)$ and $n - (p-2) \mid n - (p-2)$, thus $n - (p-2) \mid f(n) - n$ for all primes $p > 3$. Thus $f(n) - n$ has infinitely many divisors and so $f(n) = n$. It follows that the solutions of the problem are the constant functions 1, 2 and the identity function. \square

Example 5.43. Let $n > 1$ be an odd integer and let S be the set of integers $x \in \{1, 2, \dots, n\}$, such that both x and $x+1$ are relatively prime to n . Prove that

$$\prod_{x \in S} x \equiv 1 \pmod{n}.$$

Proof. Let $x \in S$, then since $\gcd(x, n) = 1$ there is a unique $y \in \{1, 2, \dots, n-1\}$ such that $xy \equiv 1 \pmod{n}$. We claim that $y \in S$. Indeed, since $n \mid xy - 1$ it is clear that $\gcd(n, y) = 1$. On the other hand, $n \mid x(y+1) - (x+1)$, thus $\gcd(n, y+1) \mid \gcd(n, x+1) = 1$ and so $\gcd(n, y+1) = 1$, proving the claim. Next, we argue as in the proof of Wilson's theorem: we create a partition of S into singletons and pairs, by putting x and y in a pair if $x \neq y$ (x, y as above) and putting x in a singleton if $x = y$. Then $\prod_{x \in S} x$ is congruent to the product of the elements of S living in singletons. These elements are those elements of S satisfying $x^2 \equiv 1 \pmod{n}$, that is $n \mid (x+1)(x-1)$. Since $\gcd(x+1, n) = 1$, we deduce that $n \mid x-1$ and so 1 is the only element of S living in a singleton. The result follows. \square

The next example is fairly challenging.

Example 5.44. (Lerch's congruence) Prove that for all odd primes p we have

$$1^{p-1} + 2^{p-1} + \dots + (p-1)^{p-1} \equiv p + (p-1)! \pmod{p^2}.$$

Proof. By Fermat's little theorem we can find integers x_1, \dots, x_{p-1} such that $j^{p-1} = 1 + px_j$ for $1 \leq j < p$. Taking the product, expanding and reducing modulo p^2 , we obtain

$$(p-1)!^{p-1} \equiv (1+px_1)(1+px_2)\dots(1+px_{p-1}) \equiv 1+p(x_1+\dots+x_{p-1}) \pmod{p^2}.$$

Next, Wilson's theorem allows us to write $(p-1)! = kp - 1$ for some integer k . Then

$$(p-1)!^{p-1} = (-1+kp)^{p-1} \equiv (-1)^{p-1} + (-1)^{p-2}(p-1)pk \equiv 1+pk \pmod{p^2}.$$

We conclude that

$$\begin{aligned} 1^{p-1} + 2^{p-1} + \dots + (p-1)^{p-1} &= p-1 + p(x_1 + \dots + x_{p-1}) \\ &\equiv p-1 + kp \equiv p + (p-1)! \pmod{p^2}, \end{aligned}$$

which is the desired result. \square

We can refine a little bit the second part of Wilson's theorem:

Proposition 5.45. *For any integer $n > 1$ the following statements are equivalent:*

- a) $n \neq 4$ and n is composite.
- b) $n \mid (n-1)!$.

Proof. Wilson's theorem easily yields that b) implies a). Assume now that a) holds and let us prove b). Write $n = ab$ with $a \geq b > 1$. If $a \neq b$ then both factors a and b appear in the product $(ab-1)! = 1 \cdot 2 \cdot \dots \cdot b \cdot \dots \cdot a \cdot \dots \cdot (ab-1)$, since $ab-1 \geq a$. Thus in this case $n = ab \mid (ab-1)! = (n-1)!$. Suppose that $a = b$, then since $n \neq 4$ we have $a > 2$. But then $ab-1 = a^2-1 > 2a$ and so the factors a and $2a$ appear in the product $(n-1)! = (a^2-1)!$, thus $n \mid 2a^2 \mid (n-1)!$ and we are done again. \square

We continue with some illustrations of the previous proposition:

Example 5.46. (Komal B 4616) For which $n > 1$ do the numbers $1!, \dots, n!$ give different remainders mod n ?

Proof. One easily checks that $n = 2, 3$ are solutions of the problem, so assume that $n > 3$ is a solution. Then precisely one of the numbers $1!, 2!, \dots, n!$ is a multiple of n and since $n \mid n!$, it follows that $(n-1)!$ is not a multiple of n . Thus by proposition 5.45 either $n = 4$ or n is a prime. One easily checks that $n = 4$ is not a solution, since $2! \equiv 3! \pmod{4}$. So n is a prime and $n \geq 5$. But Wilson's theorem yields $(n-2)! \equiv 1 = 1! \pmod{n}$, a contradiction. Hence the only solutions of the problem are 2 and 3. \square

Example 5.47. Find all positive integers n, k such that $(n-1)! + 1 = n^k$.

Proof. Note that $n > 1$ and that $n \mid (n-1)! + 1$, thus n must be prime by Wilson's theorem. One easily checks that $(n, k) = (2, 1), (3, 1), (5, 2)$ are solutions of the problem. We will prove that these are all solutions. Suppose that $n > 5$, then $n-1 > 4$ and $n-1$ is not a prime (since n is a prime), thus by proposition 5.45 we have $n-1 \mid (n-2)!$. Taking the relation $(n-2)! = n^{k-1} + n^{k-2} + \dots + n + 1$ modulo $n-1$ gives $n-1 \mid k$ and so $k \geq n-1$. But then $(n-1)! + 1 \geq n^{n-1}$ and since $(n-1)! < (n-1)^{n-1}$ we deduce that $n^{n-1} \leq (n-1)^{n-1}$, a contradiction. Hence the solutions of the problem are $(n, k) = (2, 1), (3, 1), (5, 2)$. \square

Example 5.48. Find all integers $n > 1$ for which there is a permutation a_1, a_2, \dots, a_n of $1, 2, \dots, n$ such that $\{a_1, a_1 a_2, \dots, a_1 a_2 \dots a_n\}$ is a complete residue system modulo n .

Proof. If $a_i = n$ for some $i < n$, then both $a_1 a_2 \dots a_i$ and $a_1 a_2 \dots a_{i+1}$ are multiples of n , a contradiction. Hence $a_n = n$. Then $a_1 a_2 \dots a_{n-1} = (n-1)!$ is not a multiple of n and by proposition 5.45 n is either 4 or a prime number. Conversely, if $n = 4$ we can take the permutation $a_1 = 1, a_2 = 3, a_3 = 2, a_4 = 4$, while if n is a prime number, we can consider the permutation defined by $a_1 = 1, a_n = n$ and $a_i = 1 + (i-1)^{-1}$ for $2 \leq i \leq n-1$, where $(i-1)^{-1}$ is the inverse modulo n of $i-1$, in $\{1, \dots, n-1\}$. For $2 \leq i < n$ we have

$$a_1 a_2 \dots a_i \equiv \prod_{j=2}^i j(j-1)^{-1} \equiv i \pmod{n},$$

and clearly $a_1, a_2, \dots, a_n \in \{1, 2, \dots, n\}$ are pairwise distinct, hence they form a permutation of $1, 2, \dots, n$. Therefore the answer of the problem is $n = 4$ and $n = p$ for some prime p . \square

Yet another slight but useful refinement of Wilson's theorem is the following.

Theorem 5.49. *For all primes p and all $0 \leq k \leq p-1$ we have*

$$k!(p-k-1)! + (-1)^k \equiv 0 \pmod{p}.$$

Proof. Note that $(p-1)! = k!(k+1)(k+2)\dots(p-1)$ and

$$p-1 \equiv -1 \pmod{p}, \dots, k+1 \equiv -(p-k-1) \pmod{p}.$$

Multiplying these congruences and using Wilson's theorem yields

$$-1 \equiv (p-1)! \equiv k!(-1)^{p-1-k}(p-k-1)! \pmod{p}.$$

Taking into account that $(-1)^{p-1} \equiv 1 \pmod{p}$, the result follows. \square

We continue with several illustrations of the usefulness of theorem 5.49:

Example 5.50. Prove that for all odd primes p we have

$$1!2!\dots(p-1)! \equiv (-1)^{\frac{p^2-1}{8}} \left(\frac{p-1}{2}\right)! \pmod{p}.$$

Proof. One can easily check the result for $p = 3$, so assume that $p > 3$. By theorem 5.49 we have

$$k!(p-1-k)! \equiv (-1)^{k-1} \pmod{p}$$

for $0 \leq k \leq p-1$. Taking the product for $1 \leq k \leq \frac{p-3}{2}$ yields

$$\prod_{k=1}^{\frac{p-3}{2}} k! \cdot \prod_{k=1}^{\frac{p-3}{2}} (p-1-k)! \equiv (-1)^{0+1+\dots+\frac{p-5}{2}} \pmod{p}.$$

Rearranging the factors in the left-hand side and using the identity

$$0 + 1 + \dots + \frac{p-5}{2} = \frac{p^2-1}{8} - p + 2$$

yields

$$\prod_{1 \leq k \neq \frac{p-1}{2} \leq p-2} k! \equiv (-1)^{\frac{p^2-1}{8}-p+2} \equiv -(-1)^{\frac{p^2-1}{8}} \pmod{p}.$$

Multiplying this last congruence by $\left(\frac{p-1}{2}\right)! \cdot (p-1)!$ and using Wilson's theorem finally yields the desired result. \square

Example 5.51. (China TST 2010) Prove the existence of an unbounded sequence $a_1 \leq a_2 \leq \dots$ of positive integers having the following property: for all sufficiently large integers n such that $n+1$ is composite, all prime divisors of $n!+1$ are greater than $n+a_n$.

Proof. Suppose that $p \mid n!+1$ and $n > 2$, then clearly $p > n$ since otherwise $p \mid n!$. On the other hand, by theorem 5.49 we have $(p-n-1)!n! \equiv (-1)^{n-1} \pmod{p}$ and since $n! \equiv -1 \pmod{p}$ we deduce that $(p-n-1)! \equiv (-1)^n \pmod{p}$. By assumption $n+1$ is composite so $p-n-1 > 0$. We cannot have $p-n-1 = 1$ since otherwise we would have $n = p-2$ and $1 \equiv (-1)^n \equiv$

$(-1)^{p-2} = -1 \pmod{p}$, a contradiction with $p > 2$. Hence $p - n - 1 \geq 2$ and since $(p - n - 1)! \equiv (-1)^n \pmod{p}$, we deduce that $(p - n - 1)! \geq p - 1 \geq n$. Thus, if a_n is the smallest positive integer m for which $m! \geq n$, then $p - n - 1 \geq a_n$ for all $n > 2$ and all prime factors p of $n! + 1$. It is clear that a_n is a nondecreasing unbounded sequence of positive integers. \square

Example 5.52. (JBMO TST 2013 Turkey) Find all positive integers n such that $2n + 7 \mid n! - 1$.

Proof. Since $n = 1$ is a solution, we assume in the sequel that $n > 1$. Note that if p is a prime divisor of $2n + 7$ then $p \mid n! - 1$ and so $p \geq n + 1$. If $2n + 7$ is composite, we deduce that $2n + 7 \geq (n + 1)^2$ and then $n^2 \leq 6$, forcing $n = 2$, which is not a solution of the problem.

Thus $2n + 7 = p$ is a prime and the hypothesis becomes $\left(\frac{p-7}{2}\right)! \equiv 1 \pmod{p}$. Now theorem 5.49 with $k = \frac{p-7}{2}$ combined with the previous congruence yield $\left(\frac{p+5}{2}\right)! \equiv (-1)^{\frac{p-9}{2}} \pmod{p}$. Thus

$$(-1)^{\frac{p-9}{2}} \equiv \left(\frac{p-7}{2}\right)! \prod_{j \in \{-5, -3, -1, 1, 3, 5\}} \frac{p-j}{2} \equiv \prod_{j \in \{-5, -3, -1, 1, 3, 5\}} \frac{p-j}{2} \pmod{p}.$$

Noting that $p - j \equiv -j \pmod{p}$ and simplifying the above expression, we obtain

$$64(-1)^{\frac{p-7}{2}} \equiv 15^2 = 225 \pmod{p}.$$

If $p \equiv 1 \pmod{4}$ then $p \mid 225 + 64 = 289$ thus $p = 17$, which gives the solution $n = 5$, while if $p \equiv 3 \pmod{4}$ then $p \mid 225 - 64 = 161$ which then implies $p = 23$ and $n = 8$, another solution of the problem. So 1, 5, 8 are the solutions of the problem. \square

Example 5.53. (Saint Petersburg 1996) Prove that for any prime p the numbers $1!, 2!, \dots, (p-1)!$ give at least $\lfloor \sqrt{p} \rfloor$ different remainders when divided by p .

Proof. The key idea is again the congruence

$$k!(p-1-k)! \equiv (-1)^{k-1} \pmod{p}$$

established in theorem 5.49. Multiplying it by $p - k$ yields $k!(p - k)! \equiv (-1)^k k \pmod{p}$, for $1 \leq k \leq p - 1$. Now let a_1, \dots, a_s be the distinct remainders modulo p given by the numbers $1!, 2!, \dots, (p - 1)!$. Then the previous congruence shows that each of the numbers $p - 1, 2, p - 3, 4, \dots$ is congruent to a product of two elements among a_1, \dots, a_s . There are $\frac{p-1}{2}$ different remainders mod p among $p - 1, 2, p - 3, 4, \dots$ and there are at most $\binom{s}{2} + s = \frac{s(s+1)}{2}$ possible remainders given by products of two numbers among a_1, \dots, a_s . Thus $\frac{s(s+1)}{2} \geq \frac{p-1}{2}$ and we easily deduce from this that $s \geq \lfloor \sqrt{p} \rfloor$. \square

We end this section with a beautiful but challenging problem.

Example 5.54. (IMO Shortlist 2005) Let f be a nonconstant polynomial with integer coefficients and positive leading coefficient. Prove that $f(n!)$ is composite for infinitely many integers $n \geq 1$.

Proof. Write $f(X) = a_d X^d + a_{d-1} X^{d-1} + \dots + a_0$ for some integers a_0, \dots, a_d with $a_d > 0$. If $a_0 = 0$, the result is clear, so assume that $a_0 \neq 0$. Given a prime p , the congruence $f((p - k)!) \equiv 0 \pmod{p}$ is equivalent (by theorem 5.49) to $x_k \equiv 0 \pmod{p}$, where

$$x_k = a_0(k - 1)!^d + a_1(k - 1)!^{d-1}(-1)^k + \dots + a_d(-1)^{kd}.$$

If k is large enough, say $k \geq k_0$, then $a_d^2 \mid (k - 1)!$ and $|x_k| > 2a_d^2$. Choose, for such k , a prime factor p_k of $\frac{x_k}{a_d}$. Since $\frac{x_k}{a_d} \equiv (-1)^{kd} \pmod{a_d}$, we have $\gcd(p_k, a_d) = 1$. If $p_k \leq k - 1$, then p_k divides $\frac{(k-1)!}{a_d}$, which combined with $\frac{x_k}{a_d} \equiv 0 \pmod{p_k}$ gives $p_k \mid (-1)^{kd}$, a contradiction. Thus $p_k \geq k$ for $k \geq k_0$.

Suppose now that the conclusion of the problem fails, so there is $N \geq k_0$ such that $f(n!)$ is not composite for $n \geq N$. By increasing N , we may assume that the function $x \rightarrow f(x!) - x$ is increasing and positive on $[N, \infty)$. By the previous two paragraphs we know that $p_k \geq k$ for $k \geq N$ and $p_k \mid f((p_k - k)!)$. Choose now $k = k_a = a(N + 1)! + 2$ for $a \geq 1$, so that $k, k + 1, \dots, k + N - 1$ are composite and so $p_k - k \geq N$. We conclude that $f((p_k - k)!) = p_k$ for these k . Letting $x_a = p_{k_a} - k_a$, we obtain $f(x_a!) = x_a + a(N + 1)! + 2$ for all sufficiently large a . Since the numbers (x_a) are pairwise distinct (by the previous equality), for infinitely many a we have $x_{a+1} \geq x_a + 1$ and so

$$f(x_a!) - x_a + (N + 1)! = f(x_{a+1}!) - x_{a+1} \geq f((x_a + 1)!) - (x_a + 1).$$

This implies that

$$f((x_a + 1)x_a!) - f(x_a!) \leq 1 + (N + 1)!,$$

which is impossible since $\frac{f((x_a+1)x_a!)}{f(x_a!)} \rightarrow \infty$ for $a \rightarrow \infty$. The result follows. \square

5.2.2 Application to sums of two squares

We have already seen (an easy consequence of Fermat's little theorem) that if p is a prime dividing a number of the form $x^2 + 1$ with $x \in \mathbf{Z}$, then $p = 2$ or $p \equiv 1 \pmod{4}$. The next important result establishes the converse.

Theorem 5.55. *Let p be a prime. Then the congruence $x^2 \equiv -1 \pmod{p}$ has a solution if and only if $p = 2$ or p is of the form $4k + 1$.*

Proof. We have already seen one implication, so assume that $p = 2$ or $p \equiv 1 \pmod{4}$. We need to prove the existence of an integer x such that $p \mid x^2 + 1$. If $p = 2$ pick $x = 1$, so assume that $p > 2$. Taking $k = \frac{p-1}{2}$ in theorem 5.49 and observing that k is even, we obtain

$$\left(\frac{p-1}{2}\right)!^2 \equiv -(-1)^k = -1 \pmod{p},$$

thus $x = \left(\frac{p-1}{2}\right)!$ is a solution of the congruence $x^2 \equiv -1 \pmod{p}$. \square

Remark 5.56. The proof shows that

$$\left(\frac{p-1}{2}\right)!^2 \equiv 1 \pmod{p}$$

when $p \equiv 3 \pmod{4}$, so $\left(\frac{p-1}{2}\right)! \equiv \pm 1 \pmod{p}$. Deciding for which primes p we have $\left(\frac{p-1}{2}\right)! \equiv 1 \pmod{p}$ is a rather delicate problem.

The following example is a refinement of the previous theorem.

Example 5.57. (Iran TST 2004) Let $p \equiv 1 \pmod{4}$ be a prime number. Prove that the equation $x^2 - py^2 = -1$ has solutions in positive integers.

Proof. Let (x, y) be the smallest positive solution of the Pell equation $x^2 - py^2 = 1$. Then $x^2 \equiv y^2 + 1 \pmod{4}$, which forces x being odd and y being even. Next, we have $p \mid x^2 - 1 = (x+1)(x-1)$, thus $p \mid x+1$ or $p \mid x-1$. If $p \mid x-1$, then $\frac{x-1}{2p}$ and $\frac{x+1}{2}$ are relatively prime numbers whose product is the square $(\frac{y}{2})^2$, thus $\frac{x-1}{2p} = a^2$ and $\frac{x+1}{2} = b^2$ for some positive integers a, b such that $ab = \frac{y}{2}$. Then $b^2 - pa^2 = 1$ and by minimality of the solution (x, y) we must have $a \geq y$ and so $x = 1 + 2pa^2 \geq 1 + 2py^2$, obviously impossible. Thus $p \mid x+1$ and a similar argument gives the existence of positive integers a, b such that $\frac{x+1}{2p} = a^2$ and $\frac{x-1}{2} = b^2$. Then $b^2 - pa^2 = -1$ and the result follows. \square

We can now prove the following beautiful theorem.

Theorem 5.58. (*Fermat*) Any prime $p \equiv 1 \pmod{4}$ can be written as the sum of two squares.

Proof. This follows immediately from the previous theorem and theorem 3.70. As the proof of theorem 3.70 is rather delicate, we provide now an alternative simple (but rather tricky) argument based on infinite descent. Choose an integer a such that $p \mid a^2 + 1$, which is possible by the previous theorem. Replacing a by its remainder when divided by p , we may assume that $0 < a < p$. Then $a^2 + 1 = kp$ for some positive integer k , with $k < p$.

Let r be the smallest positive integer for which rp is the sum of two squares, say $rp = x^2 + y^2$, with x, y nonnegative integers. The previous paragraph shows that $r \leq k < p$. If $r = 1$, we are done, so suppose that $r > 1$. Let x_1, y_1 be integers such that $|x_1| \leq \frac{r}{2}$, $|y_1| \leq \frac{r}{2}$ and $x \equiv x_1 \pmod{r}$, $y \equiv y_1 \pmod{r}$. Since $r \mid x^2 + y^2$, we can write $x_1^2 + y_1^2 = ru$ for some nonnegative integer u . If $u = 0$, then $r \mid \gcd(x, y)$, hence $r^2 \mid x^2 + y^2 = rp$, which is impossible, since $1 < r < p$. Thus $u > 0$. Moreover, $ru \leq 2 \cdot (r/2)^2 = r^2/2$, hence $u < r$. Finally, we have

$$r^2 up = (x^2 + y^2)(x_1^2 + y_1^2) = (xx_1 + yy_1)^2 + (xy_1 - yx_1)^2,$$

and $xx_1 + yy_1 \equiv x^2 + y^2 \equiv 0 \pmod{r}$, $xy_1 - yx_1 \equiv xy - yx \equiv 0 \pmod{r}$. Thus the previous equality exhibits up as the sum of two squares. Since $u < r$, this contradicts the minimality of r and finishes the proof. \square

We will give two more proofs of the previous theorem in the sequel. The first one uses the following very simple yet very powerful result, known as Thue's lemma.

Theorem 5.59. (*Thue's lemma*) *If a and n are relatively prime integers with $n > 1$, then there are integers x, y , not both 0, satisfying $0 \leq x, y \leq \lfloor \sqrt{n} \rfloor$ and $x \equiv \pm ay \pmod{n}$ (for a suitable choice of the sign \pm).*

Proof. Let $k = \lfloor \sqrt{n} \rfloor$, so that $k^2 \leq n < (k+1)^2$. Consider all pairs (x, y) of integers with $0 \leq x, y \leq k$. There are $(k+1)^2 > n$ such pairs, thus by the pigeonhole principle there are two different pairs (x_1, y_1) and (x_2, y_2) for which $x_1 - ay_1$ and $x_2 - ay_2$ give the same remainder when divided by n . If $x_1 = x_2$, then $ay_1 \equiv ay_2 \pmod{n}$ and so $y_1 = y_2$ since $\gcd(a, n) = 1$, a contradiction. Thus $x_1 \neq x_2$ and, by symmetry, we may assume that $x_1 < x_2$. Setting $x = x_2 - x_1$ and $y = |y_2 - y_1|$ yields the desired result. \square

Fermat's theorem 5.58 is a simple consequence of theorems 5.55 and 5.59, as follows. Let p be a prime congruent to 1 modulo 4 and pick an integer a such that $p \mid a^2 + 1$. Choose integers x, y as in Thue's lemma (theorem 5.59 above) with $n = p$. Then $x \equiv \pm ay \pmod{p}$, thus $x^2 \equiv a^2 y^2 \equiv -y^2 \pmod{p}$. It follows that $x^2 + y^2$ is a positive integer which is divisible by p and smaller than $p + p = 2p$ (since $0 \leq x, y \leq \lfloor \sqrt{p} \rfloor < \sqrt{p}$). Thus necessarily $p = x^2 + y^2$ and the result follows.

Finally, we give yet another beautiful proof of Fermat's theorem, due to Zagier. Consider a prime $p \equiv 1 \pmod{4}$ and the set

$$S = \{(x, y, z) \in \mathbf{N}^3 \mid x^2 + 4yz = p\}.$$

We will see below that we can define a map $f : S \rightarrow S$ such that $f(f(s)) = s$ for all $s \in S$ and the equation $f(x) = x$ has exactly one solution x_0 in S . It follows that $|S|$ (the number of elements of S) is odd, since we can partition S into pairs of the form $(s, f(s))$ (for $s \neq x_0$) and the singleton $\{x_0\}$. Consider now the map $g : S \rightarrow S$ sending (x, y, z) to (x, z, y) . Then clearly $g(g(s)) = s$ for all $s \in S$. If the equation $g(x) = x$ had no solution in S , then the same argument as above would imply that $|S|$ is even, a contradiction. Thus we can find $(x, y, z) \in S$ such that $g(x, y, z) = (x, y, z)$ and then $p = x^2 + 4y^2 = x^2 + (2y)^2$ is a sum of two squares.

We still need to construct the map $f : S \rightarrow S$ above. For $(x, y, z) \in S$ define $f(x, y, z)$ as follows. First, note that $x \neq y - z$ (otherwise $p = (y + z)^2$ is a perfect square, a contradiction) and $x \neq 2y$ (otherwise p is even). Next, if $x < y - z$ set $f(x, y, z) = (x + 2z, z, y - x - z)$, if $y - z < x < 2y$ set $f(x, y, z) = (2y - x, y, x - y + z)$ and finally, if $x > 2y$ set $f(x, y, z) = (x - 2y, x - y + z, y)$. A simple, yet tedious computation shows that $f(x, y, z) \in S$ and that $f(f(s)) = s$ for all $s \in S$. Moreover, the equation $f(x, y, z) = (x, y, z)$ is easily seen to have exactly one solution: for such (x, y, z) we must have $y - z < x < 2y$ and $x = y$, thus $x^2 + 4xz = p$ and then $x = 1 = y$ and $z = \frac{p-1}{4}$. The theorem is therefore proved.

Using Fermat's theorem, we can finally answer the question: which positive integers are sums of two squares? Recall that if p is a prime, then $v_p(n)$ is the exponent of the prime p in the factorization of n , i.e. the largest nonnegative integer k for which $p^k \mid n$.

Theorem 5.60. *An integer $n > 1$ is the sum of two squares if and only if $v_p(n)$ is even for all primes $p \equiv 3 \pmod{4}$ dividing n .*

Proof. Suppose that $v_p(n)$ is even for all primes $p \equiv 3 \pmod{4}$ dividing n . Thus we can write $n = 2^a \cdot m^2 \cdot p_1 \dots p_k$, where p_1, \dots, p_k are primes congruent to 1 mod 4 (not necessarily distinct) and m is a positive integer. Since 2, m^2 and each of p_1, \dots, p_k are sums of two squares (by Fermat's theorem), and since the set of sums of two squares is stable under multiplication by Lagrange's identity

$$(a^2 + b^2)(c^2 + d^2) = (ac + bd)^2 + (ad - bc)^2,$$

it follows that n is a sum of two squares.

To prove the converse, suppose that $n = a^2 + b^2$ for some integers a, b . If $p \equiv 3 \pmod{4}$ and $k = v_p(n) \geq 1$, then $p^k \mid a^2 + b^2$. By corollary 5.28, we obtain $p \mid \gcd(a, b)$. Write $a = pa_1, b = pb_1$. Then $v_p(a_1^2 + b_1^2) = k - 2$. If $k - 2 = 0$, we are done, otherwise we repeat the argument and we have $a_1 = pa_2, b = pb_2$ and $v_p(a_2^2 + b_2^2) = k - 4$. Continuing in this way we decrease k at every step by 2. At some moment we must reach 0, hence k is even. \square

Example 5.61. (USA TST 2008) Solve in integers the equation $x^2 = y^7 + 7$.

Proof. Since there are no solutions for $y < -1$, we may assume that $y + 2 > 0$. It is not difficult to see that $y \equiv 1 \pmod{4}$. We rewrite the equation as $x^2 + 11^2 = y^7 + 2^7$ or equivalently

$$x^2 + 11^2 = (y + 2)(y^6 - 2y^5 + 4y^4 - 8y^3 + 16y^2 - 32y + 64).$$

Since $y \equiv 1 \pmod{4}$, we have $y + 2 \equiv 3 \pmod{4}$, thus there exists a prime q such that $v_q(y + 2)$ is odd. Note that q does not divide $y^6 - 2y^5 + 4y^4 - 8y^3 + 16y^2 - 32y + 64$, as otherwise q would divide $7 \cdot 64$ and $x^2 + 11^2$, a contradiction. Thus $v_q(y^7 + 2^7)$ is odd, which is impossible, as it equals $v_q(x^2 + 11^2)$ and $q \equiv 3 \pmod{4}$. The result follows. \square

Example 5.62. Find the least nonnegative integer n for which there is a non-constant function $f : \mathbf{Z} \rightarrow [0, \infty)$ such that for all integers x, y

- a) $f(xy) = f(x)f(y)$;
- b) $2f(x^2 + y^2) - f(x) - f(y) \in \{0, 1, 2, \dots, n\}$.

For this n find all functions with the above properties.

Proof. Note first that for $n = 1$ there are functions satisfying a) and b). Indeed, for any prime p with $p \equiv 3 \pmod{4}$ define $f_p : \mathbf{Z} \rightarrow [0, \infty)$ by $f_p(x) = 0$ if $p|x$ and $f_p(x) = 1$, otherwise. Then a) follows from the fact that if $p|xy$ then $p|x$ or $p|y$. On the other hand $p|x^2 + y^2$ iff $p|x$ and $p|y$ (by corollary 5.28), and this implies b).

Suppose now that f is a nonconstant function that satisfies a) and b) with $n = 0$. Then $2f(x^2 + y^2) = f(x) + f(y)$ and hence

$$2f(x)^2 = 2f(x^2) = 2f(x^2 + 0) = f(x) + f(0).$$

In particular, $f(0)^2 = f(0)$. If $f(0) = 1$ then a) implies that f is the constant function 1, so $f(0) = 0$. Consequently $2f(x)^2 = f(x)$ for every $x \in \mathbf{Z}$. This together with a) imply that $f(x)^2 = f(x^2) = 2f(x^2)^2 = 2f(x)^4$. In particular, $2f(x)^2 \neq 1$ for all x and therefore f is the zero function, a contradiction. So $n = 1$ is the least integer with required properties.

We will prove now that if $n = 1$, then each nonconstant function f satisfying a) and b) is of the form f_p , or the function equal to 1 at nonzero integers

and 0 at 0. We already know that $f(0) = 0$. Since $f(1)^2 = f(1)$ and $f(1) = 0$ would make f identically zero and therefore constant, we have $f(1) = 1$. Also,

$$2f(x)^2 - f(x) = 2f(x^2 + 0) - f(x) - f(0) \in \{0, 1\}$$

for all $x \in \mathbf{Z}$, thus $f(x) \in \{0, 1\}$ for all x . (The third possibility $f(x) = \frac{1}{2}$ is excluded since it would make $f(x^2) = \frac{1}{4}$, an excluded value.) We have $f(-1)^2 = f(1) = 1$, so $f(-1) = 1$. Then $f(-x) = f(-1)f(x) = f(x)$ and it follows from a) that it suffices to find $f(p)$ for any prime p . Suppose there is $x > 0$ with $f(x) = 0$. Since $x \neq 1$ it follows that for some prime divisor p of x we have $f(p) = 0$. Suppose that there is another prime q for which $f(q) = 0$. Then $2f(p^2 + q^2) \in \{0, 1\}$ shows that $f(p^2 + q^2) = 0$. Hence for all integers a and b we have

$$0 = 2f(a^2 + b^2)f(p^2 + q^2) = 2f((ap + bq)^2 + (aq - bp)^2).$$

On the other hand $0 \leq f(x) + f(y) \leq 2f(x^2 + y^2)$ and the above identities show that $f(ap + bq) = f(aq - bp) = 0$. But p and q are relatively prime and by Bézout's lemma there are integers a and b such that $aq - bp = 1$. Then $1 = f(1) = f(aq - bp) = 0$, a contradiction. So, there is only one prime p for which $f(p) = 0$. Suppose that $p = 2$. Then $f(x) = 0$ for x even and $2f(x^2 + y^2) = 0$ for x, y odd. Hence $f(x) = f(y) = 0$ for all odd x and y , a contradiction since f is not constant. Suppose that $p \equiv 1 \pmod{4}$ and write $p = a^2 + b^2$ for some positive integers a, b (which is possible by Fermat's theorem). Then $f(a) = f(b) = 0$, but $\max(a, b) > 1$ and there is a prime q that divides it. Therefore $f(q) = 0$, a contradiction since $q < p$. Hence $p \equiv 3 \pmod{4}$ and we have that $f(x) = 0$ if x is divisible by p and $f(x) = 1$ if not. Hence $f = f_p$. \square

Example 5.63. Find all functions $f : \mathbf{N} \rightarrow \mathbf{Z}$ with the properties:

- i) $f(a) \geq f(b)$ whenever a divides b ;
- ii) $f(ab) + f(a^2 + b^2) = f(a) + f(b)$ for all $a, b \in \mathbf{N}$.

Proof. By considering the function $f(x) - f(1)$, we may assume that $f(1) = 0$, so $f(n) \leq 0$ for all n by the first condition. The second condition with $b = 1$, then reads $f(a^2 + 1) = f(1) = 0$ and in particular $f(2) = 0$.

We prove next that $f(p) = 0$ for all primes $p \equiv 1 \pmod{4}$. Indeed, take such a prime p and consider a positive integer a such that $p \mid a^2 + 1$ (it exists by theorem 5.55). Then $f(p) \geq f(a^2 + 1) = f(1) = 0$. Since $f(p) \leq 0$, we deduce that $f(p) = 0$.

Next, we observe that if $f(a) = f(b) = 0$, then $f(ab) + f(a^2 + b^2) = 0$ and $f(ab), f(a^2 + b^2) \leq 0$, hence $f(ab) = 0$. It follows immediately from this and the previous paragraph that $f(n) = 0$ whenever n is a product of primes (not necessarily distinct) congruent to 1 mod 4.

Suppose now that $\gcd(a, b) = 1$. Then $a^2 + b^2$ is a product of primes congruent to 1 mod 4, except for a possible power of 2.

Since we saw that $f(2) = 0$, the same argument as in the previous paragraph shows that $f(a^2 + b^2) = 0$ and so $f(ab) = f(a) + f(b)$.

We compute next $f(p^k)$ for a prime p . We saw that if $f(a) = f(b) = 0$ then $f(ab) = 0$, so $f(p^k) = 0$ if $p = 2$ or if $p \equiv 1 \pmod{4}$, so we may assume that $p \equiv 3 \pmod{4}$. By taking $b = a^k$ in the second relation and using that $f(a^k) \geq f(a^{k+1})$ and $f(a) \geq f(a^2 + a^{2k})$, we deduce that both of these inequalities are equalities and so $f(a^k) = f(a^{k+1})$ for all a and k . We conclude that $f(p^k) = f(p)$.

Putting everything together we deduce that if $n = p_1^{k_1} \dots p_r^{k_r}$ for some distinct primes p_1, \dots, p_r and k_1, \dots, k_r positive integers, then $f(n) = f(p_1) + \dots + f(p_r)$ and each $f(p_i)$ is 0 if $p_i = 2$ or $p_i \equiv 1 \pmod{4}$. This determines f uniquely if we fix the values of $f(p)$ for all primes $p \equiv 3 \pmod{4}$. This gives us a family of solutions and we will check now that we can allow arbitrary values at these primes.

So, choose any function g defined on the set of primes $p \equiv 3 \pmod{4}$ and define $f(1) = f(2) = 0$ and $f(p) = g(p)$ if $p \equiv 3 \pmod{4}$, $f(p) = 0$ for the other primes p and extend f to all positive integers by

$$f(p_1^{k_1} \dots p_r^{k_r}) = f(p_1) + \dots + f(p_r).$$

We have to check that f is a solution. But the first relation is clear and the second one follows by considering the prime factorization of $a, b, \gcd(a, b)$ and using the fact that for $\gcd(a, b) = 1$ the prime factors of $a^2 + b^2$ are all congruent 2 or 1 (mod 4), on which f vanishes. \square

5.3 Lagrange's theorem and applications

5.3.1 The number of solutions of polynomial congruences

Fermat's little theorem has the striking consequence that for any prime p the polynomial $X^p - X$ has p different zeros modulo p , namely $0, 1, \dots, p-1$. There is another polynomial having such zeros, namely $X(X-1)\dots(X-p+1)$. Of course, $X^p - X$ and $X(X-1)\dots(X-p+1)$ are not equal as polynomials. In this section we will define a congruence relation for polynomials with integer coefficients and we will prove that $X^p - X$ and $X(X-1)\dots(X-p+1)$ are congruent modulo p . Using this, we will study the map $x \mapsto x^d \pmod{p}$ when d is a positive integer and p is a prime. This study will play a key role in the last chapter.

Let us start by introducing a congruence relation between polynomials. We denote by $\mathbf{Z}[X]$ the set of polynomials with integer coefficients. The following definition should not be a great surprise for the reader.

Definition 5.64. Let n be an integer and let $f, g \in \mathbf{Z}[X]$. We say that f and g are congruent modulo n and write $f \equiv g \pmod{n}$ if all coefficients of the polynomial $f - g$ are multiples of n , in other words, if there is $h \in \mathbf{Z}[X]$ such that $f - g = nh$.

We note straight away one common mistake: if $f \equiv g \pmod{n}$ then clearly $f(x) \equiv g(x) \pmod{n}$ for all integers x . However, the converse **does not** hold: take $f = X^2 + X$ and $g = 2$, then $f(x) \equiv g(x) \equiv 0 \pmod{2}$ for all integers x , however f is not congruent to g modulo 2, since the coefficients of $X^2 + X - 2$ are not all even.

As an example, the polynomials $X(X-1)(X-2)$ and $X^3 - X$ are congruent modulo 3 since the coefficients of their difference

$$(X^3 - X) - X(X-1)(X-2) = 3X(X-1)$$

are multiples of 3. On the other hand, $X^3 - X$ and $X(X-1)(X-2)$ are not congruent modulo n for any $n > 1$ different from 3.

Just as for integers, one can immediately prove the following formal properties of congruences for polynomials. We leave the simple proofs to the reader.

Proposition 5.65. For all polynomials $f, g, h, k \in \mathbf{Z}[X]$ and all n we have

- a) $f \equiv f \pmod{n}$.
- b) If $f \equiv g \pmod{n}$, then $g \equiv f \pmod{n}$.
- c) If $f \equiv g \pmod{n}$ and $g \equiv h \pmod{n}$, then $f \equiv h \pmod{n}$.
- d) If $f \equiv g \pmod{n}$ and $h \equiv k \pmod{n}$, then $f + h \equiv g + k \pmod{n}$ and $fh \equiv gk \pmod{n}$.

Example 5.66. Prove that for all $f, g \in \mathbf{Z}[X]$ and all primes p we have

$$(f + g)^p \equiv f^p + g^p \pmod{p} \quad \text{and} \quad f(X)^p \equiv f(X^p) \pmod{p}.$$

Proof. The first congruence follows directly from the binomial formula

$$(f + g)^p = f^p + g^p + \sum_{k=1}^{p-1} \binom{p}{k} f^{p-k} g^k$$

and the fact that $p \mid \binom{p}{k}$ for $1 \leq k \leq p-1$. For the second congruence, write $f(X) = a_0 + a_1X + \dots + a_nX^n$. Applying repeatedly the first congruence yields

$$f(X)^p = (a_0 + a_1X + \dots + a_nX^n)^p \equiv a_0^p + (a_1X)^p + \dots + (a_nX^n)^p \pmod{p}.$$

Using Fermat's little theorem we obtain $a_i^p \equiv a_i \pmod{p}$, and the result follows. \square

The next very useful result extends the usual property of primes (if p divides ab then p divides a or b) to polynomials.

Theorem 5.67. (*Gauss' lemma for polynomials*) Let p be a prime and let f, g be polynomials with integer coefficients such that $f \cdot g \equiv 0 \pmod{p}$. Then $f \equiv 0 \pmod{p}$ or $g \equiv 0 \pmod{p}$.

Proof. Assume that this is not the case and write

$$f(X) = a_0 + a_1X + \dots + a_dX^d, \quad g = b_0 + b_1X + \dots + b_eX^e$$

for some integers $a_0, \dots, a_d, b_0, \dots, b_e$. Let i be the smallest nonnegative integer for which p does not divide a_i (i exists since by assumption f is not congruent to 0 modulo p). Similarly, let j be the smallest nonnegative integer for which

p does not divide b_j . The coefficient of X^{i+j} in $f(X)g(X)$ is $\sum_{u+v=i+j} a_u b_v$ and by assumption it is divisible by p . On the other hand, if $u+v=i+j$ and $(u, v) \neq (i, j)$, then $u < i$ or $v < j$, thus $a_u b_v$ is divisible by p . It follows that

$$0 \equiv \sum_{u+v=i+j} a_u b_v \equiv a_i b_j \pmod{p},$$

which contradicts the fact that a_i and b_j are not divisible by p . The result follows. \square

The fundamental link between congruences of polynomials and solutions of polynomial congruences is the following

Theorem 5.68. *Let a be an integer and let $f \in \mathbf{Z}[X]$. Then $f(a) \equiv 0 \pmod{n}$ if and only if there is $g \in \mathbf{Z}[X]$ such that $f(X) \equiv (X - a)g(X) \pmod{n}$. Moreover, if this is the case then we can choose g of degree less than or equal to $\deg(f) - 1$.*

Proof. Suppose first that such g exists. By definition there is a polynomial h with integer coefficients such that $f(X) = (X - a)g(X) + nh(X)$. Plugging in $X = a$ yields $f(a) = nh(a) \equiv 0 \pmod{n}$. Suppose conversely that $f(a) \equiv 0 \pmod{n}$. Write $f(X) = c_0 + c_1X + \dots + c_dX^d$ for some integers c_0, \dots, c_d and note that

$$f(X) - f(a) = c_1(X - a) + c_2(X^2 - a^2) + \dots + c_d(X^d - a^d) = (X - a)g(X),$$

with

$$g(X) = c_1 + c_2(X + a) + \dots + c_d(X^{d-1} + \dots + a^{d-1}),$$

a polynomial with integer coefficients of degree less than or equal to $d - 1$. Since $f(X) - (X - a)g(X) = f(a)$ and $f(a) \equiv 0 \pmod{n}$, we have $f(X) \equiv (X - a)g(X) \pmod{n}$ and we are done. \square

We can establish now the following very important result, which is the mod p analogue of the fact that any nonzero polynomial f with complex coefficients has at most $\deg f$ distinct roots.

Theorem 5.69. (Lagrange) *Let p be a prime and let f be a polynomial with integer coefficients. If at least one of the coefficients of f is not a multiple of p (in other words if f is not congruent to 0 mod p), then the congruence $f(x) \equiv 0 \pmod{p}$ has at most $\deg f$ solutions.*

Proof. We prove this by induction on the degree d of f . The case $d = 0$ being clear, assume that the result holds for d and let us prove it for $d + 1$. Let $f \in \mathbf{Z}[X]$ be a polynomial of degree $d + 1$ which is not congruent to 0 mod p . If the congruence $f(x) \equiv 0 \pmod{p}$ has no solutions, we are done, so assume that this is not the case and pick a solution a . The previous theorem shows the existence of a polynomial $g \in \mathbf{Z}[X]$ such that $f(X) \equiv (X - a)g(X) \pmod{p}$ and $\deg(g) \leq d$. Note that g is not 0 mod p , since f is not 0 mod p . Thus by the inductive hypothesis the congruence $g(x) \equiv 0 \pmod{p}$ has at most d solutions. Since each solution of the congruence $f(x) \equiv 0 \pmod{p}$ is either a or a solution of the congruence $g(x) \equiv 0 \pmod{p}$ (this crucially uses the fact that p is a prime, contrary to all previous arguments), the result follows. \square

Remark 5.70. The result is completely false for congruences $f(x) \equiv 0 \pmod{n}$, where n is composite. For instance the congruence $x^3 \equiv x \pmod{6}$ has 6 solutions, yet the polynomial $X^3 - X$ is certainly not congruent to 0 mod 6.

The following very useful result is an immediate consequence of Fermat's little theorem and Lagrange's theorem.

Theorem 5.71. *For all primes p we have*

$$X^{p-1} - 1 \equiv (X - 1)(X - 2)\dots(X - p + 1) \pmod{p}.$$

Proof. Let f be the difference between the left-hand side and the right-hand side. Then $\deg f \leq p - 2$, since $X^{p-1} - 1$ and $(X - 1)\dots(X - p + 1)$ are monic of degree $p - 1$. On the other hand Fermat's little theorem yields $f(i) \equiv 0 \pmod{p}$ for $1 \leq i \leq p - 1$, hence by Lagrange's theorem $f \equiv 0 \pmod{p}$, as desired. \square

The previous theorem encodes a large family of congruences, among which is Wilson's theorem $(p - 1)! + 1 \equiv 0 \pmod{p}$. Indeed, this follows by looking

at the constant terms of the polynomials appearing in the previous theorem. By looking at the coefficient of X^{p-1-i} with $1 \leq i < p-1$, we obtain

$$\sum_{1 \leq k_1 < k_2 < \dots < k_i < p} k_1 k_2 \dots k_i \equiv 0 \pmod{p}.$$

The following rather interesting examples illustrate the power of the previous theorems.

Example 5.72. (Romania TST 2001) Find all pairs (m, n) of positive integers, with $m, n \geq 2$, such that $a^n - 1$ is divisible by m for each $a \in \{1, 2, 3, \dots, n\}$.

Proof. Let p be a prime factor of m , so that $p \mid a^n - 1$ for $1 \leq a \leq n$. If $p \leq n$, we obtain $p \mid p^n - 1$, a contradiction. Thus $p \geq n+1$. It follows that $1, 2, \dots, n$ are pairwise distinct solutions of the polynomial congruence $x^n \equiv 1 \pmod{p}$. Thus the polynomial congruence

$$x^n - 1 - (x-1)\dots(x-n) \equiv 0 \pmod{p}$$

has degree at most $n-1$ and at least n different solutions. Lagrange's theorem implies that

$$X^n - 1 \equiv (X-1)(X-2)\dots(X-n) \pmod{p}.$$

Considering the coefficients of X^{n-1} , we deduce that $p \mid \frac{n(n+1)}{2}$. Since $p > n$, the only possibility is $p = n+1$. In particular, $n+1$ is a prime $p > 2$ and m has a unique prime factor, namely p . We will show that p^2 cannot divide $a^{p-1} - 1$ for all $1 \leq a \leq p-1$, establishing therefore that $m = p$. Indeed, note that

$$(p-1)^{p-1} - 1 \equiv (-1)^{p-1} + (-1)^{p-2}(p-1)p - 1 \equiv -p(p-1) \pmod{p^2}$$

and so p^2 does not divide $(p-1)^{p-1} - 1$. □

Example 5.73. (Iran TST 2011) Let p be a prime, k a positive integer and let $f \in \mathbf{Z}[X]$ such that p^k divides $f(x)$ for all $x \in \mathbf{Z}$. If $k \leq p$, prove that there are polynomials $g_0, g_1, \dots, g_k \in \mathbf{Z}[X]$ such that

$$f(X) = \sum_{i=0}^k p^{k-i} (X^p - X)^i \cdot g_i(X).$$

Proof. We will prove this result by induction on k . Suppose first that $k = 1$ and write $f(X) = (X^p - X)q(X) + r(X)$ for some polynomials $q, r \in \mathbf{Z}[X]$ such that $\deg r < p$ (this is possible since $X^p - X$ is monic). The hypothesis combined with Fermat's little theorem show that $p \mid r(x)$ for all integers x . Since $\deg r < p$, Lagrange's theorem yields $r \equiv 0 \pmod{p}$ and the result follows.

Let us prove the inductive step. Assume that the result holds for k , that $k + 1 \leq p$ and that p^{k+1} divides $f(x)$ for all x . By the inductive hypothesis there are polynomials $g_i \in \mathbf{Z}[X]$ such that

$$f(X) = \sum_{i=0}^k p^{k-i} (X^p - X)^i \cdot g_i(X).$$

If x and z are any integers and if $y = \frac{x^p - x}{p}$ (an integer by Fermat's little theorem), the binomial formula gives

$$(x + pz)^p - (x + pz) \equiv p(y - z) \pmod{p^2},$$

therefore

$$f(x + pz) \equiv \sum_{i=0}^k p^k (y - z)^i g_i(x + pz) \equiv p^k \sum_{i=0}^k (y - z)^i g_i(x) \pmod{p^{k+1}}.$$

We conclude that p divides $\sum_{i=0}^k (y - z)^i g_i(x)$ for any x and z , and replacing z with $y - z$, it follows that $\sum_{i=0}^k z^i g_i(x) \equiv 0 \pmod{p}$ for all integers z and x . Since $k < p$, Lagrange's theorem yields $g_i(x) \equiv 0 \pmod{p}$ for all i and all x . Applying the base case, we can find $h_i, r_i \in \mathbf{Z}[X]$ such that

$$g_i(X) = (X^p - X)h_i(X) + pr_i(X).$$

Replacing these expressions in $f(X) = \sum_{i=0}^k p^{k-i} (X^p - X)^i \cdot g_i(X)$ finishes the inductive step. \square

Example 5.74. (USA TST 2009) Let $p \geq 5$ be a prime and let a, b, c be integers such that p does not divide $(a - b)(b - c)(c - a)$. Let $i, j, k \geq 0$ be integers such that $p - 1 \mid i + j + k$ and such that for all integers x

$$p \mid (x - a)(x - b)(x - c)[(x - a)^i (x - b)^j (x - c)^k - 1].$$

Prove that the numbers i, j, k are divisible by $p - 1$.

Proof. Using Fermat's little theorem, we may replace i, j, k with their remainders mod $p - 1$, without affecting the hypothesis or the conclusion. Thus we may assume that $0 \leq i, j, k < p - 1$ and need to prove that $i = j = k = 0$. Assume that this is not the case. Since $p - 1 \mid i + j + k$, we deduce that $i + j + k = p - 1$ or $2(p - 1)$. If $i + j + k = 2(p - 1)$, we replace each $x \in \{i, j, k\}$ with $p - 1 - x$, which does not change the hypothesis or the conclusion. Thus we may assume that $i + j + k = p - 1$. Finally, we may assume that $i = \max(i, j, k)$.

Multiplying the congruence

$$(x - a)(x - b)(x - c)[(x - a)^i(x - b)^j(x - c)^k - 1] \equiv 0 \pmod{p}$$

by $(x - a)^{j+k}$ and using Fermat's little theorem, we obtain

$$f(x) := (x - a)(x - b)(x - c)[(x - b)^j(x - c)^k - (x - a)^{j+k}] \equiv 0 \pmod{p}.$$

for all x . Since $p \geq 5$, we have

$$\deg(f) \leq 3 + j + k - 1 \leq 2 + \frac{2(p-1)}{3} < p$$

and so Lagrange's theorem yields $f(X) \equiv 0 \pmod{p}$. Combining this with theorem 5.67, we obtain

$$(X - b)^j(X - c)^k \equiv (X - a)^{j+k} \pmod{p}.$$

Since $i < p - 1$ and $i + j + k = p - 1$, we have $j + k \neq 0$, thus $(X - b)^j(X - c)^k$ vanishes at b or c . We deduce that p divides $(b - a)^{j+k}$ or $(c - a)^{j+k}$, which contradicts the hypothesis. Thus $i = j = k = 0$ and the result follows. \square

Example 5.75. (China TST 2009) Prove the existence of a number $c > 0$ such that for any prime p there are at most $cp^{2/3}$ positive integers n for which p divides $n! + 1$.

Proof. Let $p > 2$ be a prime and let $1 < n_1 < n_2 < \dots < n_m < p$ be all solutions of the congruence $n! \equiv -1 \pmod{p}$ (note that if $p \mid n! + 1$ then

$n < p$). We may assume that $m > 1$, otherwise we are done. Combining the congruences $n_i! \equiv -1 \pmod{p}$ and $n_{i+1}! \equiv -1 \pmod{p}$ yields

$$(n_i + 1)(n_i + 2) \dots (n_i + n_{i+1} - n_i) \equiv 1 \pmod{p}.$$

Lagrange's theorem shows that for each $1 \leq k < p$ the congruence

$$(x + 1)(x + 2) \dots (x + k) \equiv 1 \pmod{p}$$

has at most k solutions. We deduce that for each $1 \leq k < p$ there are at most k indices i such that $n_{i+1} - n_i = k$. This is the key point of the proof, the remaining part of the argument being purely combinatorial.

Choose a positive integer j such that

$$\frac{(j+1)(j+2)}{2} \geq m \geq \frac{j(j+1)}{2}.$$

Since for any $k \in \{1, 2, \dots, p-1\}$ the equation $n_{i+1} - n_i = k$ has at most k solutions i and since $m \geq \frac{j(j+1)}{2} = \sum_{i=1}^j j$, we deduce that when the differences $n_{i+1} - n_i$ are written in ascending order, the first is at least 1, the next two are at least 2, and so on, each time the next i differences are at least i . It follows that

$$\sum_{i=1}^{m-1} (n_{i+1} - n_i) \geq 1^2 + 2^2 + \dots + j^2 = \frac{j(j+1)(2j+1)}{6}$$

and so

$$p > n_m - n_1 \geq \frac{j(j+1)(2j+1)}{6}.$$

In particular, $p > \frac{j^3}{3}$ and $j < (3p)^{1/3}$. Since $m \leq (j+1)^2 \leq 4j^2$, the result follows. \square

5.3.2 The congruence $x^d \equiv 1 \pmod{p}$

After this series of examples, we come back to more theoretical issues. An immediate consequence of Lagrange's theorem is the following innocent-looking but nontrivial result.

Corollary 5.76. *Let p be a prime and let k be a positive integer such that $x^k \equiv 1 \pmod{p}$ for all integers x which are not multiples of p . Then $p-1 \mid k$.*

Proof. Let $d = \gcd(k, p-1)$, then $d \mid p-1$ and moreover for all x not divisible by p we have $x^d \equiv 1 \pmod{p}$ (since $x^k \equiv 1 \pmod{p}$ by assumption and $x^{p-1} \equiv 1 \pmod{p}$ by Fermat's little theorem). Thus the congruence $x^d \equiv 1 \pmod{p}$ has at least $p-1$ solutions. Lagrange's theorem yields $d \geq p-1$. Since $d = \gcd(k, p-1)$, the result follows. \square

We obtain now immediately the following very important and useful congruence (which is not very easy to prove directly).

Corollary 5.77. *a) If j is a positive integer, not divisible by $p-1$, then*

$$1^j + 2^j + \dots + (p-1)^j \equiv 0 \pmod{p}.$$

b) If f is a polynomial with integer coefficients and $\deg(f) < p-1$, then

$$f(0) + f(1) + \dots + f(p-1) \equiv 0 \pmod{p}.$$

Proof. a) By the previous corollary we can choose an integer x which is not divisible by p and such that p does not divide $x^j - 1$. Let $S = 1^j + 2^j + \dots + (p-1)^j$. Since the remainders of $x, 2x, \dots, (p-1)x$ when divided by p are a permutation of $1, 2, \dots, p-1$, we obtain

$$x^j S = x^j + (2x)^j + \dots + ((p-1)x)^j \equiv 1^j + 2^j + \dots + (p-1)^j \equiv S \pmod{p},$$

thus p divides $S(x^j - 1)$. Since p does not divide $x^j - 1$, the result follows.

b) Write $f(X) = a_0 + a_1 X + \dots + a_d X^d$ for some integers a_0, \dots, a_d and $d < p-1$. Then

$$f(0) + f(1) + \dots + f(p-1) = pa_0 + a_1(1+2+\dots+(p-1)) + \dots + a_d(1^d + \dots + (p-1)^d).$$

By part a) each of the sums $1+2+\dots+(p-1), \dots, 1^d + \dots + (p-1)^d$ is divisible by p . The result follows. \square

Before illustrating the previous results with some concrete examples, we would like to discuss in more detail the congruence $x^d \equiv 1 \pmod{p}$ where d is a positive integer and p is a prime. This will play a crucial role in chapter 6. For this note that we can always reduce the study to the case $d \mid p-1$, since the congruence $x^d \equiv 1 \pmod{p}$ has exactly the same solutions as the congruence $x^{\gcd(d, p-1)} \equiv 1 \pmod{p}$ (by Fermat's little theorem and the fact that $\gcd(x^d - 1, x^{p-1} - 1) = x^{\gcd(d, p-1)} - 1$). Again, Fermat's little theorem combined with Lagrange's theorem easily yield the following result.

Theorem 5.78. *Let p be a prime and let d be a positive divisor of $p-1$. Then the congruence $x^d \equiv 1 \pmod{p}$ has exactly d solutions.*

Proof. Since $d \mid p-1$, we can find a polynomial with integer coefficients $f(X)$ such that $X^{p-1} - 1 = (X^d - 1)f(X)$ (explicitly, $f(X) = 1 + X^d + \dots + X^{(\frac{p-1}{d}-1)d}$). By Fermat's little theorem the congruence $x^{p-1} \equiv 1 \pmod{p}$ has $p-1$ solutions. Each solution of this congruence is a solution of one of the congruences $x^d \equiv 1 \pmod{p}$ and $f(x) \equiv 0 \pmod{p}$. By Lagrange's theorem, these two congruences have at most d , respectively $p-1-d$ solutions. Since in total they have $p-1 = d + p-1-d$ solutions, we deduce that the first one has d solutions and the second one $p-1-d$ solutions. The result follows. \square

Let us illustrate the previous results with some concrete examples.

Example 5.79. A Carmichael number is a positive integer n such that $n \mid a^n - a$ for any integer a .

a) Prove that n is a Carmichael number if and only if n is squarefree and $p-1$ divides $n-1$ for any prime p dividing n .

b) Find all Carmichael numbers of the form $3pq$ with p, q primes.

Proof. a) Suppose that n is a Carmichael number, then n divides $p^n - p$ for any prime p . Thus if $p \mid n$, p^2 cannot divide n (otherwise we would obtain $p^2 \mid p^n - p$ and then $p^2 \mid p$). Thus n is squarefree. Next, if $p \mid n$ is a prime then $p \mid a^{n-1} - 1$ for any a relatively prime to p and so $p-1 \mid n-1$ by corollary 5.76. The converse follows from example 5.3.

b) By part a) we obtain that $3, p, q$ are distinct and that $p-1 \mid 3pq-1$ and $q-1 \mid 3pq-1$. The first congruence implies that $p-1 \mid 3q-1$, while the second

yields $q - 1 \mid 3p - 1$. We may assume that $p > q$, so that $3q - 1 < 3(p - 1)$. Thus either $p - 1 = 3q - 1$ (impossible, as $p \neq 3$) or $2(p - 1) = 3q - 1$. So $2p = 3q + 1$ and since $q - 1 \mid 3p + 1$, we immediately obtain that $q - 1 \mid 9q + 1$. This forces $q - 1 \mid 10$ and we easily infer that $q = 11$ and $p = 17$. Thus $n = 561$ is the only Carmichael number of the form $3pq$. \square

Example 5.80. (Romania TST 2008) Let n be an integer greater than 1. Compute the greatest common divisor of the numbers $2^n - 2, 3^n - 3, \dots, n^n - n$ for given n .

Proof. For $n = 2$ the answer is 2, so assume that $n > 2$. Let

$$d = \gcd(2^n - 2, \dots, n^n - n)$$

and let p be a prime factor of d . If $p > n$, then the congruence of degree n $x^n \equiv x \pmod{p}$ has pairwise distinct solutions $0, 1, \dots, n$ modulo p , a contradiction with Lagrange's theorem. Thus $p \leq n$. In particular $d \mid p^n - p$ and so p^2 cannot divide d . Next, $p \mid a^{n-1} - 1$ for all a relatively prime to p , since $p \mid a^n - a$ for $1 \leq a \leq n$ and $n \geq p$. Corollary 5.76 gives $p - 1 \mid n - 1$. Conversely, if p is a prime such that $p - 1 \mid n - 1$ then $p \mid a^n - a$ for all integers a and so $p \mid d$. In other words, we have just proved that

$$d = \prod_{p-1 \mid n-1} p. \quad \square$$

Example 5.81. (IMO 1997 Shortlist) Let p be a prime and let f be a polynomial with integer coefficients such that $f(0) = 0$, $f(1) = 1$ and $f(x)$ is congruent to 0 or 1 modulo p for all integers x . Prove that $\deg(f) \geq p - 1$.

Proof. Assuming the contrary, corollary 5.77 yields

$$f(0) + f(1) + \dots + f(p-1) \equiv 0 \pmod{p}.$$

But the left-hand side is congruent to a sum of zeros and ones by assumption, and there is at least one zero and at least one 1 in this sum. It is thus impossible to get a multiple of p . \square

Example 5.82. (Mathematical Reflections O 21) Find the least degree of a nonconstant polynomial f with integer coefficients having the property that $f(0), f(1), \dots, f(p-1)$ are all perfect $(p-1)$ th powers.

Proof. Let f be such a polynomial and write $f(i) = x_i^{p-1}$ for some integers x_0, \dots, x_{p-1} . By Fermat's little theorem we deduce that $f(i)$ is congruent to 0 or 1 mod p for all $0 \leq i \leq p-1$. Assume that $\deg f < p-1$, then corollary 5.77 gives

$$f(0) + f(1) + \dots + f(p-1) \equiv 0 \pmod{p}$$

and since each of the numbers $f(0), \dots, f(p-1)$ is congruent to 0 or 1 mod p we deduce that $f(0), \dots, f(p-1)$ are all congruent to 0 mod p or all congruent to 1 mod p . Thus there is $\varepsilon \in \{0, 1\}$ such that the congruence $f(x) \equiv \varepsilon \pmod{p}$ has at least p solutions, which contradicts Lagrange's theorem. Thus $\deg f \geq p-1$. Since $f(X) = X^{p-1}$ obviously satisfies the required properties, we conclude that the answer is $p-1$. \square

Example 5.83. (Giuga) Let n be an integer greater than 1. Prove that

$$n \mid 1 + 1^{n-1} + 2^{n-1} + \dots + (n-1)^{n-1}$$

if and only if for every prime divisor p of n ,

$$p \mid \frac{n}{p} - 1 \quad \text{and} \quad p-1 \mid \frac{n}{p} - 1$$

Proof. Let p be a prime divisor of n . Let us see when p divides $1 + S$, where $S = 1^{n-1} + 2^{n-1} + \dots + (n-1)^{n-1}$. Write $n = kp$ for a positive integer k . Then each nonzero remainder modulo p appears exactly k times among $1, 2, \dots, n-1$, hence

$$1 + S \equiv 1 + k(1^{n-1} + 2^{n-1} + \dots + (p-1)^{n-1}).$$

By corollary 5.77 the number $1^{n-1} + 2^{n-1} + \dots + (p-1)^{n-1}$ is congruent to 0 modulo p if $p-1$ does not divide $n-1$, and it is congruent to -1 modulo p otherwise. We conclude that $p \mid 1 + S$ if and only if $p-1$ divides $n-1$ (equivalent to $p-1 \mid \frac{n}{p} - 1$) and $p \mid k-1 = \frac{n}{p} - 1$.

This already proves one implication: if n divides $1 + S$, then $p-1 \mid n-1$ and $p \mid \frac{n}{p} - 1$ for all $p \mid n$. Conversely, suppose that these conditions are

satisfied. Since $p \mid \frac{n}{p} - 1$ for all $p \mid n$, it follows that n is squarefree. Hence n divides $1 + S$ if and only if $p \mid 1 + S$ for any $p \mid n$. By the first paragraph, this is true, which concludes the proof. \square

Remark 5.84. Giuga's conjecture is that the only numbers satisfying the previous divisibility are the prime numbers. Note that the condition $p - 1 \mid \frac{n}{p} - 1$ is equivalent to $p - 1 \mid n - 1$, in other words any number satisfying the divisibility is a Carmichael number. Let us call $n > 1$ a Giuga number if n is composite and $p \mid \frac{n}{p} - 1$ for all prime divisors p of n (which implies that n is squarefree). We can rephrase Giuga's conjecture as saying that no Giuga number is also a Carmichael number. The first Giuga numbers are

$$30, 858 = 2 \cdot 3 \cdot 11 \cdot 13, 1722 = 2 \cdot 3 \cdot 7 \cdot 41, \dots$$

and there are also monster Giuga numbers such as

$$2 \cdot 3 \cdot 11 \cdot 23 \cdot 31 \cdot 47059 \cdot 2259696349 \cdot 110725121051.$$

It is not known if there are infinitely many Giuga numbers. An excellent exercise for the reader is to check the equivalence of the following statements:

- a) n is a Giuga number;
- b) $1^{\varphi(n)} + 2^{\varphi(n)} + \dots + (n-1)^{\varphi(n)} \equiv -1 \pmod{n}$;
- c) $\sum_{p \mid n} \frac{1}{p} - \prod_{p \mid n} \frac{1}{p}$ is a positive integer.

A beautiful exposition of these results (and many others) can be found in the article "Giuga's conjecture on primality", by D. Borwein, J. M. Borwein, P. B. Borwein and R. Girgensohn, published in the American Mathematical Monthly, vol. 103, No 1, 1996.

We give now a more conceptual proof of example 5.44, based on corollary 5.77.

Example 5.85. (Lerch's congruence) Prove that for all odd primes p we have

$$1^{p-1} + 2^{p-1} + \dots + (p-1)^{p-1} \equiv p + (p-1)! \pmod{p^2}.$$

Proof. Let us write

$$f(X) = \prod_{i=1}^{p-1} (X - i) = X^{p-1} + a_{p-2}X^{p-2} + \dots + a_1X + a_0$$

for some integers a_0, \dots, a_{p-2} . Since by theorem 5.71

$$\prod_{i=1}^{p-1} (X - i) \equiv X^{p-1} - 1 \pmod{p},$$

we have $p \mid a_1, \dots, a_{p-2}$ and $a_0 = (p-1)!$. Next observe that

$$0 = \sum_{i=1}^{p-1} f(i) = \sum_{i=1}^{p-1} i^{p-1} + \sum_{j=0}^{p-2} a_j (1^j + 2^j + \dots + (p-1)^j).$$

Since $1^j + 2^j + \dots + (p-1)^j \equiv 0 \pmod{p}$ for $1 \leq j \leq p-2$ (by corollary 5.77), all terms $a_j(1^j + 2^j + \dots + (p-1)^j)$ with $1 \leq j \leq p-2$ are multiples of p^2 . It follows that

$$1^{p-1} + 2^{p-1} + \dots + (p-1)^{p-1} \equiv -(p-1)(p-1)! \pmod{p^2}.$$

It suffices therefore to prove that

$$-(p-1)(p-1)! \equiv p + (p-1)! \pmod{p^2},$$

which reduces to $(p-1)! \equiv -1 \pmod{p}$, i.e. Wilson's theorem. \square

5.3.3 The Chevalley-Warning theorem

We will prove now a stunningly beautiful result about the number of solutions of some systems of polynomial congruences, known as the Chevalley-Warning theorem. This will require the next result, which is a simple but rather powerful multi-variable version of corollary 5.77.

Corollary 5.86. *Let $F \in \mathbf{Z}[X_1, \dots, X_n]$ be a polynomial with integer coefficients in the variables X_1, \dots, X_n and let p be a prime such that $\deg F < n(p-1)$. Then*

$$\sum_{(x_1, \dots, x_n) \in \{0, 1, \dots, p-1\}^n} F(x_1, \dots, x_n) \equiv 0 \pmod{p}.$$

Proof. The polynomial F is a linear combination with integer coefficients of monomials of the form $X_1^{i_1} \dots X_n^{i_n}$ with $i_1 + \dots + i_n < n(p-1)$, since $\deg F < n(p-1)$ by assumption. Thus it suffices to prove the result for each such monomial, i.e. that

$$\sum_{(x_1, \dots, x_n) \in \{0, 1, \dots, p-1\}^n} x_1^{i_1} \dots x_n^{i_n} \equiv 0 \pmod{p}$$

whenever i_1, \dots, i_n are nonnegative integers with $i_1 + \dots + i_n < n(p-1)$. Since

$$\sum_{(x_1, \dots, x_n) \in \{0, 1, \dots, p-1\}^n} x_1^{i_1} \dots x_n^{i_n} = \left(\sum_{x_1=0}^{p-1} x_1^{i_1} \right) \cdot \dots \cdot \left(\sum_{x_n=0}^{p-1} x_n^{i_n} \right),$$

it is enough to prove that $p \mid \sum_{x=0}^{p-1} x^{i_j}$ for some $j \in \{1, 2, \dots, n\}$. But since $i_1 + \dots + i_n < n(p-1)$, there is some j for which $i_j < p-1$ and for this j we have $p \mid \sum_{x=0}^{p-1} x^{i_j}$ by corollary 5.77. \square

We are now ready to prove the following result, which was conjectured by Artin.

Theorem 5.87. (*Chevalley-Warning*) *Let p be a prime and let k and n be positive integers. Let f_1, \dots, f_k be polynomials with integer coefficients in the variables X_1, \dots, X_n , such that*

$$n > \sum_{i=1}^k \deg f_i.$$

Then the number of n -tuples $(x_1, \dots, x_n) \in \{0, 1, \dots, p-1\}^n$ such that

$$f_1(x_1, \dots, x_n) \equiv f_2(x_1, \dots, x_n) \equiv \dots \equiv f_k(x_1, \dots, x_n) \equiv 0 \pmod{p}$$

is a multiple of p .

Proof. The following proof is rather magical. Consider the polynomial

$$F = (1 - f_1^{p-1})(1 - f_2^{p-1}) \dots (1 - f_k^{p-1})$$

and note that by assumption $\deg F < (p-1)n$. The key observation is that for any $x = (x_1, \dots, x_n) \in \{0, 1, \dots, p-1\}^n$ the simultaneous congruences

$$f_1(x) \equiv f_2(x) \equiv \dots \equiv f_k(x) \equiv 0 \pmod{p}$$

are equivalent to the single congruence $F(x) \equiv 1 \pmod{p}$. Indeed, by Fermat's little theorem $f_i(x)^{p-1} \equiv 1 \pmod{p}$ unless $f_i(x) \equiv 0 \pmod{p}$, thus $F(x) \equiv 0 \pmod{p}$ unless $f_i(x) \equiv 0 \pmod{p}$ for all $1 \leq i \leq k$.

Now, let N be the number of n -tuples $(x_1, \dots, x_n) \in \{0, 1, \dots, p-1\}^n$ such that $F(x_1, \dots, x_n) \equiv 1 \pmod{p}$. Then clearly

$$\sum_{(x_1, \dots, x_n) \in \{0, 1, \dots, p-1\}^n} F(x_1, \dots, x_n) \equiv N \pmod{p},$$

thus it suffices to prove that the left-hand side is a multiple of p . But this is the content of corollary 5.86. \square

A very useful (yet straightforward) consequence of the Chevalley-Waring theorem is the following result, which guarantees the existence of nontrivial solutions to systems of polynomial congruences, as long as these systems have enough unknowns and a trivial solution.

Corollary 5.88. *Under the assumptions of the Chevalley-Waring theorem, if $f_i(0, \dots, 0) = 0$ for all i then the system*

$$f_1(x_1, \dots, x_n) \equiv f_2(x_1, \dots, x_n) \equiv \dots \equiv f_k(x_1, \dots, x_n) \equiv 0 \pmod{p}$$

has a solution (x_1, \dots, x_n) with at least one x_i not divisible by p .

Proof. The Chevalley-Waring theorem says that the number of solutions of the system is divisible by p . The assumption that $f_i(0, \dots, 0) = 0$ ensures that $(0, 0, \dots, 0)$ is a solution of the system. It follows that the system has a solution different from this one, which finishes the proof. \square

Example 5.89. Let p be a prime and let a, b, c be integers. Prove that there are integers x, y, z , not all divisible by p , such that $p \mid ax^2 + by^2 + cz^2$.

Proof. This is an immediate consequence of corollary 5.88. \square

We have already proved the result below in example 4.39, but the proof given there was not very natural. We give now a very conceptual proof based on the Chevalley-Warning theorem (more precisely on corollary 5.88).

Example 5.90. (Erdős-Ginzburg-Ziv) Let p be a prime. Prove that among any $2p - 1$ integers there are p whose sum is a multiple of p .

Proof. Applying corollary 5.88 to

$$f_1(X) = \sum_{i=1}^{2p-1} a_i X_i^{p-1}, \quad f_2(X) = \sum_{i=1}^{2p-1} X_i^{p-1}$$

yields the existence of $(x_1, \dots, x_{2p-1}) \in \{0, 1, \dots, p-1\}^{2p-1}$ such that not all x_i 's are multiples of p and

$$f_1(x_1, \dots, x_{2p-1}) \equiv f_2(x_1, \dots, x_{2p-1}) \equiv 0 \pmod{p}.$$

Choosing $I = \{i \mid x_i \not\equiv 0 \pmod{p}\}$, Fermat's little theorem yields

$$\sum_{i \in I} a_i \equiv 0 \pmod{p}, \quad \sum_{i \in I} 1 \equiv 0 \pmod{p}.$$

The second congruence and the inequalities $1 \leq |I| \leq 2p - 1$ yield $|I| = p$. Thus $(a_i)_{i \in I}$ satisfy all requirements. \square

Remark 5.91. The result still holds without the assumption that p is a prime, but the case of primes is the most difficult. See the reduction to the case of a prime given in the proof of example 4.39.

Example 5.92. (Zimmerman) a) Let p be a prime and let a_1, \dots, a_{2p-1} be integers. If I is a subset of $\{1, \dots, 2p-1\}$ with p elements, let $S_I = \sum_{i \in I} a_i$. Prove that

$$\sum_I S_I^{p-1} \equiv 0 \pmod{p},$$

the sum being taken over all subsets I with p elements of $\{1, 2, \dots, 2p-1\}$.

b) Deduce a new proof of the Erdős-Ginzburg-Ziv theorem.

Proof. a) Let S be the left-hand side. Brutally expanding each S_I^{p-1} , we see that we can write

$$S = \sum_{\substack{k_1, \dots, k_{2p-1} \geq 0 \\ k_1 + \dots + k_{2p-1} = p-1}} c_{k_1, \dots, k_{2p-1}} a_1^{k_1} \dots a_{2p-1}^{k_{2p-1}}$$

for some integers $c_{k_1, \dots, k_{2p-1}}$. Let us fix a monomial $a_1^{k_1} \dots a_{2p-1}^{k_{2p-1}}$ and analyze which subsets I contribute to this monomial. Note that at most $p-1$ of the k_i 's are positive, say precisely j of them are positive. Now I contributes to this monomial if and only if it contains all the positive k_i , and all such I have the same contribution. There are $\binom{2p-1-j}{p-j}$ sets I with p elements, containing the positive k_i 's. Note that this last binomial coefficient is a multiple of p (for instance by Lucas' theorem). It follows that the coefficient of each $a_1^{k_1} \dots a_{2p-1}^{k_{2p-1}}$ is a multiple of p , and the result follows.

b) Let a_1, \dots, a_{2p-1} be integers and use the notations of the previous exercise. We need to prove that some S_I is a multiple of p . Assuming that this is not the case, it follows from Fermat's little theorem and the previous exercise that

$$\binom{2p-1}{p} \equiv 0 \pmod{p}.$$

This is absurd, since $\binom{2p-1}{p} \mid (p+1)(p+2)\dots(p+p-1)$ and so it is not a multiple of p . \square

We end this section with a more challenging application of the Chevalley-Waring theorem.

Example 5.93. (IMO Shortlist 2003) Let p be a prime number and let A be a set of positive integers such that:

a) the set of prime divisors of the elements of A consists of $p-1$ elements and

b) for any nonempty subset of A , the product of its elements is not a perfect p -th power.

What is the largest possible number of elements of A ?

Proof. It is not difficult to see that A can have $(p-1)^2$ elements: pick pairwise distinct primes q_1, \dots, q_{p-1} and let the elements of A be

$$q_1, q_1^{1+p}, \dots, q_1^{1+p(p-2)}, \dots, q_{p-1}, q_{p-1}^{1+p}, \dots, q_{p-1}^{1+p(p-2)}.$$

Clearly A has $(p-1)^2$ elements and satisfies a). To see that A satisfies b), pick a nonempty subset B of A and choose a prime factor q_j of $\prod_{x \in B} x$. Suppose that $q_j^{1+px_1}, \dots, q_j^{1+px_k}$ are all elements of B that are divisible by q_j , then the exponent of q_j in the prime factorization of $\prod_{x \in B} x$ is

$$v_{q_j}(\prod_{x \in B} x) = k + p(x_1 + \dots + x_k)$$

and this is clearly not divisible by p since $1 \leq k \leq p-1$. Thus $\prod_{x \in B} x$ is not a perfect p th power.

We move now to the difficult part of the problem, namely proving that any such set A has at most $(p-1)^2$ elements. Suppose that a set A satisfying a) and b) has more than $(p-1)^2$ elements, and choose $k = (p-1)^2 + 1$ pairwise distinct elements x_1, \dots, x_k of A . Let q_1, \dots, q_{p-1} be the different prime divisors of $\prod_{x \in A} x$. Write for $1 \leq j \leq k$

$$x_j = q_1^{e_{1j}} q_2^{e_{2j}} \dots q_{p-1}^{e_{p-1j}}$$

for some integers $e_{i,j}$ and consider the polynomials

$$f_i(X_1, \dots, X_k) = X_1^{p-1} e_{i1} + X_2^{p-1} e_{i2} + \dots + X_k^{p-1} e_{ik}$$

for $1 \leq i \leq p-1$. Then

$$\sum_{i=1}^{p-1} \deg f_i = (p-1)^2 < k,$$

thus by corollary 5.88 the system

$$f_1(z_1, \dots, z_k) \equiv \dots \equiv f_{p-1}(z_1, \dots, z_k) \equiv 0 \pmod{p}$$

has a nontrivial solution $(z_1, \dots, z_k) \in \{0, 1, \dots, p-1\}^k$. Letting

$$I = \{i \in \{1, \dots, k\} | z_i \neq 0\},$$

Fermat's little theorem yields

$$\sum_{j \in I} e_{ij} \equiv 0 \pmod{p}$$

for all $1 \leq i \leq p-1$. It follows that $\prod_{j \in I} x_j$ is a perfect p th power, contradicting the fact that A satisfies b). Thus the answer of the problem is $(p-1)^2$. \square

5.4 Quadratic residues and quadratic reciprocity

We now turn to the study of the congruence $x^2 \equiv a \pmod{p}$, where p is a prime and a is an integer. The case $p = 2$ being clear (in this case $x^2 \equiv x \pmod{p}$ for all x , thus the congruence has exactly one solution, $x \equiv a \pmod{p}$), we will assume in this whole section that $p > 2$. We therefore fix an odd prime p in the sequel.

5.4.1 Quadratic residues and Legendre's symbol

Let us introduce the following useful terminology.

Definition 5.94. If a is an integer, we say that a is a quadratic residue mod p if the congruence $x^2 \equiv a \pmod{p}$ has solutions. Otherwise, we say that a is a quadratic non-residue mod p . We say that a residue class \bar{a} is a quadratic residue class if a is a quadratic residue mod p (or equivalently if any integer in the residue class is a quadratic residue mod p).

Since $x^2 \equiv y^2 \pmod{p}$ if and only if $x \equiv \pm y \pmod{p}$, it is clear that the quadratic residues in $\{0, 1, \dots, p-1\}$ are precisely those of $0^2, 1^2, \dots, (\frac{p-1}{2})^2$, and these are pairwise distinct, so there are $\frac{p+1}{2}$ quadratic residue classes mod p , and $\frac{p-1}{2}$ nonzero quadratic residue classes mod p . Since this is extremely useful in practice, let us glorify this result:

Proposition 5.95. *For each odd prime p there are exactly $\frac{p+1}{2}$ quadratic residues mod p (and thus $\frac{p-1}{2}$ nonzero quadratic residues mod p), and these are the residues of $0^2, 1^2, \dots, (\frac{p-1}{2})^2$.*

Example 5.96. Prove that if a, b, c are integers such that p does not divide abc , then the congruence $ax^2 + by^2 \equiv c \pmod{p}$ has at least one solution.

Proof. Let A be the set of remainders mod p of the numbers ax^2 when $0 \leq x \leq \frac{p-1}{2}$ and similarly let B be the set of remainders mod p of the numbers $c - by^2$ when $0 \leq y \leq \frac{p-1}{2}$. Then A and B consist each of $\frac{p+1}{2}$ distinct remainders mod p (since p does not divide ab and the numbers x^2 with $0 \leq x \leq \frac{p-1}{2}$ are pairwise distinct modulo p). Since $|A| + |B| > p$, we deduce that $A \cap B \neq \emptyset$, which is exactly the desired statement. \square

We introduce now a very useful and important arithmetic function, Legendre's symbol. Much of this section is devoted to the study of the basic properties of this function.

Definition 5.97. (Legendre's symbol) Let a be an integer and let p be an odd prime. We define $\left(\frac{a}{p}\right) = 0$ if $p \mid a$, $\left(\frac{a}{p}\right) = 1$ if a is a nonzero quadratic residue mod p and $\left(\frac{a}{p}\right) = -1$ otherwise.

So we obtain a map

$$\left(\frac{\cdot}{p}\right) : \mathbf{Z} \rightarrow \{-1, 0, 1\}$$

called Legendre's symbol mod p . This map enjoys a certain number of remarkable properties. The first property is its p -periodicity, i.e.

$$\left(\frac{a + kp}{p}\right) = \left(\frac{a}{p}\right)$$

for all integers a and all k . This is immediate from the definition.

In order to establish the second important property of Legendre's symbol, we will need the following analogue of theorem 5.71.

Theorem 5.98. For all odd primes p we have

$$X^{\frac{p-1}{2}} - 1 \equiv \prod_{i=1}^{\frac{p-1}{2}} (X - i^2) \pmod{p}.$$

Proof. The proof is very similar to that of theorem 5.71: the difference between the two sides is a polynomial of degree at most $\frac{p-1}{2} - 1$ whose values at $1^2, 2^2, \dots, \left(\frac{p-1}{2}\right)^2$ are divisible by p (since $(i^2)^{\frac{p-1}{2}} = i^{p-1} \equiv 1 \pmod{p}$ for $1 \leq i \leq \frac{p-1}{2}$ by Fermat's little theorem). Lagrange's theorem combined with the fact that $1^2, 2^2, \dots, \left(\frac{p-1}{2}\right)^2$ are pairwise distinct modulo p yield the desired result.

Note the following alternate and simpler argument: letting $f(X)$ be the difference between the left-hand side and the right-hand side, we obtain

$$f(X^2) = X^{p-1} - 1 - \prod_{i=1}^{\frac{p-1}{2}} (X^2 - i^2) \equiv X^{p-1} - 1 - \prod_{i=1}^{p-1} (X - i) \equiv 0 \pmod{p},$$

the last congruence being a consequence of theorem 5.71. The result follows immediately. \square

We are now ready to prove the following beautiful:

Theorem 5.99. (*Euler's criterion*) For all a and all odd primes $p > 2$ we have

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \pmod{p}.$$

In particular, if a is not divisible by p , then a is a quadratic residue mod p , i.e. $\left(\frac{a}{p}\right) = 1$ if and only if $a^{\frac{p-1}{2}} \equiv 1 \pmod{p}$.

Proof. The result is clear when a is a multiple of p , so assume that this is not the case. Note that $(a^{\frac{p-1}{2}})^2 \equiv 1 \pmod{p}$ by Fermat's little theorem, therefore $a^{\frac{p-1}{2}} \equiv \pm 1 \pmod{p}$. From theorem 5.98 with $X = a$, we see that $a^{\frac{p-1}{2}} \equiv 1 \pmod{p}$ if and only if a is a quadratic residue modulo p . \square

A very useful consequence of the previous theorem is the following result, that we have actually already encountered when discussing Fermat's little theorem (see corollary 5.28 for instance).

Corollary 5.100. *For all odd primes p we have*

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}},$$

so -1 is a quadratic residue mod p if and only if $p \equiv 1 \pmod{4}$.

The previous theorem also implies the very important:

Theorem 5.101. *For all integers a, b we have*

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right) \cdot \left(\frac{b}{p}\right).$$

Proof. By Euler's criterion, both sides are congruent to $(ab)^{\frac{p-1}{2}}$ modulo p , in particular the difference between the left-hand side and the right-hand side is a multiple of p . But since this difference is a number between -2 and 2 , and since $p > 2$, this difference must be 0 . \square

Note that the only nontrivial statement in the previous theorem is the rather surprising fact that if a, b are quadratic non-residues mod p , then their product ab is a quadratic residue mod p . We illustrate now the previous results with many examples.

Example 5.102. Let p be an odd prime. Find all functions $f : \mathbf{Z} \rightarrow \mathbf{Z}$ such that for all integers m, n we have

- a) if p divides $m - n$ then $f(m) = f(n)$;
- b) $f(mn) = f(m)f(n)$.

Proof. Clearly the constant functions 0 and 1 are solutions of the problem, so suppose from now on that f is not constant. Since f is multiplicative and nonconstant we have $f(1) = 1$. Then for all n not divisible by p we have (by Fermat's little theorem) $1 = f(1) = f(n^{p-1}) = f(n)^{p-1}$, thus $f(n) = \pm 1$ for such n . Also, note that $f(0) = f(n)f(0)$ for all n , thus $f(0) = 0$ and so $f(n) = 0$ whenever n is divisible by p . Next, note that if x is a quadratic residue mod p and not divisible by p , then $f(x) = 1$ (write $x \equiv y^2 \pmod{p}$ with y not divisible by p , then $f(x) = f(y^2) = f(y)^2 = 1$). Choose n not divisible by p such that $\left(\frac{n}{p}\right) = -1$. If x runs over the nonzero quadratic

residues mod p , then nx runs over all quadratic non-residues mod p , and $f(nx) = f(n)f(x) = f(n)$. Thus f is constant on quadratic non-residues mod p , and this constant is 1 or -1 . We conclude that there are four solutions to our problem: $f \equiv 1$, $f \equiv 0$, $f(n) = 1$ for n not divisible by p and $f(n) = 0$ for n divisible by p , and finally the Legendre symbol mod p . \square

The next example is fairly interesting: it gives an example of a polynomial f with integer coefficients which has no rational root and yet which has roots modulo any prime number, i.e. such that the congruence $f(x) \equiv 0 \pmod{p}$ has solutions for any prime p .

Example 5.103. Let p be a prime. Prove that the congruence $x^8 \equiv 16 \pmod{p}$ has at least one solution.

Proof. The key observation is the factorization

$$x^8 - 16 = (x^4 - 4)(x^4 + 4) = (x^2 - 2)(x^2 + 2)((x - 1)^2 + 1)((x + 1)^2 + 1).$$

Thus we have to prove that at least one of the congruences

$$\begin{aligned} x^2 &\equiv 2 \pmod{p}, & x^2 &\equiv -2 \pmod{p}, \\ (x - 1)^2 + 1 &\equiv 0 \pmod{p}, & (x + 1)^2 &\equiv -1 \pmod{p} \end{aligned}$$

has a solution. This is clear for $p = 2$, so assume that $p > 2$. Then we need to show that at least one of $-1, 2, -2$ is a quadratic residue mod p . But if -1 and 2 are quadratic non-residues, then their product -2 is a quadratic residue and we are done. \square

Example 5.104. Prove that if $p > 2$, then the least (positive) quadratic non-residue mod p is less than $\frac{1}{2} + \sqrt{p}$.

Proof. Let n be the smallest positive quadratic non-residue mod p . Write $p = qn + r$ with $0 \leq r < n$ and note that clearly $r > 0$, so $\left(\frac{n-r}{p}\right) = 1$ (by minimality of n). Since $n - r \equiv (q + 1)n \pmod{p}$, we have

$$1 = \left(\frac{n-r}{p}\right) = \left(\frac{(q+1)n}{p}\right) = \left(\frac{q+1}{n}\right) \cdot (-1),$$

thus $q + 1$ is a quadratic non-residue mod p . We deduce that $q + 1 \geq n$, thus $p \geq n(n + 1) + 1$, which immediately yields the desired estimate. \square

Example 5.105. a) Prove that if $p > 3$, then the sum of the quadratic residues mod p in $\{0, 1, \dots, p-1\}$ is a multiple of p .

b) Prove that if $p \equiv 1 \pmod{4}$, then the sum of quadratic residues mod p in $\{0, 1, \dots, p-1\}$ is $\frac{p(p-1)}{4}$.

Proof. a) This follows immediately from theorem 5.98 or by using the fact that the quadratic residues mod p in $\{0, 1, \dots, p-1\}$ are the remainders mod p of $0, 1^2, \dots, \left(\frac{p-1}{2}\right)^2$, thus their sum is congruent mod p to

$$1^2 + 2^2 + \dots + \left(\frac{p-1}{2}\right)^2 = \frac{p(p^2-1)}{24} \equiv 0 \pmod{p},$$

the last congruence being clear since $p > 3$ (thus $24 \mid p^2 - 1$).

b) Suppose that $p \equiv 1 \pmod{4}$. Then for all k , we have that k is a quadratic residue mod p if and only if $p-k$ is a quadratic residue mod p (since -1 is a quadratic residue mod p). Therefore we can create a partition of the set of quadratic residues mod p in $\{1, 2, \dots, p-1\}$ in classes with two elements, the sum of the elements in each class being p . Since there are $\frac{p-1}{2}$ quadratic residues between 1 and $p-1$, there will be $\frac{p-1}{4}$ such classes and so the total sum of quadratic residues is $\frac{p-1}{4} \cdot p = \frac{p(p-1)}{4}$. \square

Example 5.106. Let p be a prime of the form $4k+3$ and let m be the number of quadratic residues mod p between $\frac{p}{2}$ and p (excluding p). Prove that

$$\left(\frac{p-1}{2}\right)! \equiv (-1)^m \pmod{p}.$$

Proof. Let $a = \left(\frac{p-1}{2}\right)!$. A classical consequence of Wilson's theorem (see theorem 5.55 and the remark following it) gives $a^2 \equiv 1 \pmod{p}$, thus $a \equiv \pm 1 \pmod{p}$. In particular we have $a \equiv \left(\frac{a}{p}\right) \pmod{p}$. On the other hand we have

$$\left(\frac{a}{p}\right) = \prod_{k=1}^{\frac{p-1}{2}} \left(\frac{k}{p}\right).$$

In the above product, we can restrict ourselves to those k between 1 and $\frac{p-1}{2}$ which are quadratic non-residues (as when k is a quadratic residue the corresponding factor $\left(\frac{k}{p}\right)$ equals 1). Now, note that since $p \equiv 3 \pmod{4}$ we have $\left(\frac{-1}{p}\right) = -1$, thus an integer a is a quadratic residue if and only if $p-a$ is a quadratic non-residue. We deduce that the number of quadratic non-residues between 1 and $\frac{p-1}{2}$ is equal to the number of quadratic residues between $\frac{p}{2}$ and p (the map $x \mapsto p-x$ establishing a bijection between the corresponding sets), and this is m by definition. We conclude that

$$\left(\frac{a}{p}\right) = \prod_{k=1}^{\frac{p-1}{2}} \left(\frac{k}{p}\right) = (-1)^m,$$

which finishes the proof. \square

Example 5.107. Let p be a prime number of the form $4k+1$. Prove that

$$\sum_{j=1}^{\frac{p-1}{4}} \lfloor \sqrt{jp} \rfloor = \frac{p^2-1}{12}.$$

Proof. Write $p = 4k+1$ and observe that

$$\sum_{j=1}^k \lfloor \sqrt{jp} \rfloor = \sum_{j=1}^k \sum_{i^2 \leq jp} 1 = \sum_{i=1}^{2k} \sum_{k \geq j \geq \frac{i^2}{p}} 1.$$

As $\frac{i^2}{p}$ is not an integer, the inequality $j \geq \frac{i^2}{p}$ is equivalent to $j \geq 1 + \left[\frac{i^2}{p}\right]$. Thus we can also write

$$\sum_{j=1}^k \lfloor \sqrt{jp} \rfloor = \sum_{i=1}^{2k} \left(k - \left[\frac{i^2}{p}\right] \right) = 2k^2 - \sum_{i=1}^{2k} \left[\frac{i^2}{p}\right]$$

and the problem is reduced to

$$\sum_{i=1}^{2k} \left[\frac{i^2}{p}\right] = \frac{2k^2-2k}{3}.$$

Since the remainder of i^2 when divided by p is $i^2 - p \left\lfloor \frac{i^2}{p} \right\rfloor$ and since

$$\sum_{i=1}^{2k} i^2 = \frac{pk(2k+1)}{3},$$

we only need to prove that the sum of the quadratic residues mod p is pk , which has already been established in example 5.105. \square

We end this section with a very beautiful and challenging problem.

Example 5.108. (USA TST 2014) Find all functions $f : \mathbf{N} \rightarrow \mathbf{Z}$ such that $(m-n)(f(m)-f(n))$ is a perfect square for all m, n .

Proof. Clearly any function f of the form $f(x) = a^2x + b$ with a, b integers is a solution of the problem. We will prove that these are the only solutions. Let f be a solution of the problem and assume without loss of generality that f is not constant. Note that since $f(n+1) - f(n)$ is a perfect square for all n , the number $\gcd(f(2) - f(1), f(3) - f(2), \dots)$ is a perfect square, say a^2 , with a a positive integer. Since a^2 divides $f(n+1) - f(n)$ for all n , an immediate induction it divides $f(n) - f(1)$ for all n . Also, the function $g(x) = \frac{f(x) - f(1)}{a^2}$ still has the property that $(m-n)(g(m)-g(n))$ is a perfect square for all m, n , and moreover $\gcd(g(2) - g(1), g(3) - g(2), \dots) = 1$. Thus replacing f with g , we may assume that $a = 1$, i.e. that $\gcd(f(2) - f(1), f(3) - f(2), \dots) = 1$. We will prove that $f(n+1) - f(n) = 1$ for all n , which will finish the proof.

Suppose that there is n such that $f(n+1) - f(n)$ is a perfect square greater than 1, and fix a prime factor p of $f(n+1) - f(n)$. Let r be the remainder of $f(n)$ when divided by p and let S be the set of solutions of the congruence $f(x) \equiv r \pmod{p}$ (thinking of S as a set of residue classes rather than a set of integers in the following), thus $n, n+1 \in S$.

Now let x be the smallest quadratic non-residue in $\{2, 3, \dots, p-1\}$, so that $x-1$ is a quadratic residue mod p . If $a, b \in S$, we claim that $(1-x)a + xb = a + x(b-a) \in S$. This is clear if $a = b$, so assume that $a \neq b$ and let $m = a + x(b-a)$. We need to prove that $f(m) \equiv f(a) \pmod{p}$. Assume that this is not the case and let $c = (b-a)(f(m)-f(a))$, thus c is nonzero mod p . On the other hand by assumption $(m-a)(f(m)-f(a))$ and $(m-b)(f(m)-f(b))$ are perfect squares, thus xc and $(x-1)(b-a)(f(m)-f(b))$ are perfect squares

and in particular quadratic residues mod p . Note that $(b-a)(f(m)-f(b)) \equiv c \pmod{p}$ (as $f(a) \equiv f(b) \equiv \pmod{p}$), thus xc and $(x-1)c$ are quadratic residues mod p , while x is a quadratic non-residue and $x-1$ is a quadratic residue. This is obviously impossible, proving that $c \equiv 0 \pmod{p}$, as needed.

Now let $T = \{s-n | s \in S\}$, thus $0, 1 \in T$ (since $n, n+1 \in S$) and, thanks to the previous paragraph, $xa + (1-x)b \in T$ whenever $a, b \in T$. In particular $xT \subset T$ and $(1-x)T \subset T$. We deduce that for all $a \in T$ we have

$$a + 1 \equiv x \cdot x^{p-2}a + (1-x) \cdot (1-x)^{p-2} \cdot 1 \in T,$$

and since $0 \in T$, it immediately follows that T contains all residue classes and therefore S contains all residue classes. We deduce that $p \mid f(n) - r$ for all n , thus $p \mid f(n+1) - f(n)$ for all n , a contradiction with

$$\gcd(f(2) - f(1), f(3) - f(2), \dots) = 1. \quad \square$$

5.4.2 Points on spheres mod p and Gauss sums

Let us come back for a while to our original goal: discuss the congruence $x^2 \equiv a \pmod{p}$. If a is a multiple of p , the congruence has only one solution $x \equiv 0 \pmod{p}$, so assume that a is not a multiple of p . If x and y are solutions of the congruence then $x^2 \equiv a \equiv y^2 \pmod{p}$, thus p divides $x^2 - y^2 = (x+y)(x-y)$ and so $y \equiv \pm x \pmod{p}$. It follows that the congruence has exactly two solutions: if x is a solution, then all solutions are x and $-x$ (note that x and $-x$ are different modulo p , since $p > 2$ and a is not divisible by p). To summarize, the congruence has two solutions when $\left(\frac{a}{p}\right) = 1$ and zero solutions when $\left(\frac{a}{p}\right) = -1$. In other words, we have just obtained the following result.

Proposition 5.109. *If a is an integer and $p > 2$ is a prime, then the congruence $x^2 \equiv a \pmod{p}$ has exactly $1 + \left(\frac{a}{p}\right)$ solutions.*

The previous proposition is very useful when computing sums related to Legendre's symbol. Let us give one very important example. Consider an integer a and the congruence $x^2 - y^2 \equiv a \pmod{p}$ (in two variables x, y). If $a \equiv 0 \pmod{p}$, this is equivalent to $(x-y)(x+y) \equiv 0 \pmod{p}$ and the

solutions are given by (x, x) and $(x, -x)$ for $x \in \{0, 1, \dots, p-1\}$. Note that the solution $(0, 0)$ is counted twice, so we obtain $2p-1$ solutions. Consider now the case $a \neq 0$. Then the congruence is equivalent to $(x-y)(x+y) \equiv a \pmod{p}$. The substitution $x+y=u$, $x-y=v$ realizes a bijection between solutions of this congruence and solutions of the congruence $uv \equiv a \pmod{p}$ (note that we can recover uniquely x, y from u, v thanks to the fact that p is odd). On the other hand, if $uv \equiv a \pmod{p}$, then u and v are nonzero mod p and for each nonzero $u \pmod{p}$ there is a unique $v \pmod{p}$ such that $uv \equiv a \pmod{p}$. Thus the congruence $uv \equiv a \pmod{p}$ has $p-1$ solutions. To summarize, the congruence

$$x^2 - y^2 \equiv a \pmod{p}$$

has $p-1$ solutions when a is not a multiple of p , and $2p-1$ solutions otherwise. Let us count now the solutions in a different way. Namely, fix y and consider the congruence $x^2 \equiv y^2 + a \pmod{p}$. By the previous proposition, this congruence has $1 + \left(\frac{y^2+a}{p}\right)$ solutions. Varying y , we deduce that the total number of solutions is

$$p + \sum_{y=0}^{p-1} \left(\frac{y^2+a}{p}\right).$$

Comparing the two expressions for the number of solutions, we deduce the following result.

Proposition 5.110. *For an integer a we have*

$$\sum_{k=0}^{p-1} \left(\frac{a+k^2}{p}\right) = p-1 \quad \text{if } p \mid a \quad \text{and} \quad \sum_{k=0}^{p-1} \left(\frac{a+k^2}{p}\right) = -1 \quad \text{otherwise.}$$

The following result is a simple consequence of the previous one, and we leave the proof to the reader.

Proposition 5.111. *Let a, b, c be integers such that p does not divide a . Then*

$$\sum_{k=0}^{p-1} \left(\frac{ak^2 + bk + c}{p}\right) = (p-1) \left(\frac{a}{p}\right) \quad \text{if } p \mid b^2 - 4ac$$

and

$$\sum_{k=0}^{p-1} \left(\frac{ak^2 + bk + c}{p} \right) = - \left(\frac{a}{p} \right) \quad \text{otherwise.}$$

In particular, for any integers a, b which are not congruent mod p we have

$$\sum_{k=0}^{p-1} \left(\frac{(k+a)(k+b)}{p} \right) = -1.$$

We can use proposition 5.110 to give a very simple proof of the following beautiful result, which is not very simple to prove directly, since $x^2 + y^2$ has no simple factorization, contrary to $x^2 - y^2$.

Proposition 5.112. *The number of solutions of the congruence $x^2 + y^2 \equiv a \pmod{p}$ is $p + (p-1)(-1)^{\frac{p-1}{2}}$ if $p \mid a$ and $p - (-1)^{\frac{p-1}{2}}$ otherwise.*

Proof. Fixing y , the congruence $x^2 \equiv a - y^2 \pmod{p}$ has exactly $1 + \left(\frac{a-y^2}{p} \right)$ solutions, thus, by varying y , the total number of solutions of the congruence $x^2 + y^2 \equiv a \pmod{p}$ is

$$p + \sum_{y=0}^{p-1} \left(\frac{a - y^2}{p} \right).$$

On the other hand

$$\sum_{y=0}^{p-1} \left(\frac{a - y^2}{p} \right) = \sum_{y=0}^{p-1} \left(\frac{-1}{p} \right) \cdot \left(\frac{y^2 - a}{p} \right) = \left(\frac{-1}{p} \right) \cdot \sum_{y=0}^{p-1} \left(\frac{y^2 - a}{p} \right).$$

Since the previous proposition gives us the value of $\sum_{y=0}^{p-1} \left(\frac{y^2 - a}{p} \right)$ and since $\left(\frac{-1}{p} \right) = (-1)^{\frac{p-1}{2}}$, the result follows by combining the previous observations. \square

Before moving on, we give some concrete and quite beautiful applications of the previous proposition.

Example 5.113. Given an odd prime p , prove that the congruence

$$x^2 + y^2 + z^2 \equiv 0 \pmod{p}$$

has exactly p^2 solutions.

Proof. Fixing z , the number of solutions of the congruence $x^2 + y^2 \equiv -z^2 \pmod{p}$ is given by the previous proposition: this number is $p + (p-1)(-1)^{\frac{p-1}{2}}$ when $p \mid z$ and $p - (-1)^{\frac{p-1}{2}}$ otherwise. Since there are $p-1$ nonzero possible z , we obtain that the total number of solutions is

$$p + (p-1)(-1)^{\frac{p-1}{2}} + (p-1)(p - (-1)^{\frac{p-1}{2}}) = p^2. \quad \square$$

Example 5.114. (Iran 2015) Let $p > 5$ be a prime. Prove that at least one of the numbers $1 + p, 1 + 2p, 1 + 3p, \dots, 1 + (p-3)p$ is the sum of squares of two integers.

Proof. Suppose that the congruence $x^2 + y^2 \equiv 1 \pmod{p}$ has a nontrivial solution (x, y) , i.e. a solution with xy not divisible by p . Since $(\pm x, \pm y)$ is also a solution of the congruence, we may assume that $0 < x, y \leq \frac{p-1}{2}$. Therefore

$$1 + p \leq x^2 + y^2 \leq \frac{(p-1)^2}{2} \leq 1 + (p-3)p,$$

the last inequality being immediate for $p \geq 5$. Therefore the problem is solved if we prove the existence of such a solution. This is immediate if we prove that the congruence $x^2 + y^2 \equiv 1 \pmod{p}$ has at least 5 solutions (since there are only 4 trivial solutions). But proposition 5.112 shows that this congruence has either $p+1$ or $p-1$ solutions. Thus, as long as $p-1 \geq 5$, we are done. \square

Example 5.115. (Bulgaria TST 2007) Let p be a prime of the form $4k+3$. Consider all numbers of the form $(x^2 + y^2)^2$ with x and y integers not divisible by p . Find the number of different remainders these numbers give when divided by p .

Proof. Clearly any such remainder is a quadratic residue mod p . Since $p \equiv 3 \pmod{4}$, 0 is not among these remainders (for if $p \mid (x^2 + y^2)^2$ then $p \mid x^2 + y^2$, thus $p \mid x$ and $p \mid y$, a contradiction). Conversely, we will prove that any nonzero quadratic residue mod p appears among these remainders. It suffices to prove that for any a not divisible by p one of the congruences $x^2 + y^2 \equiv a \pmod{p}$ and $x^2 + y^2 \equiv -a \pmod{p}$ has solutions with x, y not divisible by p . Since -1 is not a quadratic residue mod p , one of the numbers a and $-a$ is not a quadratic residue mod p , say it is a . We know that the number of solutions

of the congruence $x^2 + y^2 \equiv a \pmod{p}$ is $p - (-1)^{\frac{p-1}{2}} = p + 1$, by proposition 5.112. For any such solution x and y are not divisible by p (for if $p \mid x$, then $y^2 \equiv a \pmod{p}$), contradicting the fact that a is not a quadratic residue mod p). The claim is thus proved. It follows that there are exactly $\frac{p-1}{2}$ remainders mod p . \square

Example 5.116. (USA TST 2016) Is there a nonconstant polynomial f with integer coefficients such that for all $n > 2$ the numbers $f(0), f(1), \dots, f(n-1)$ give at most $0.499n$ different remainders when divided by n ?

Proof. We will prove that there is such a polynomial. First of all, note that it suffices to check that $f(0), f(1), \dots, f(n-1)$ give at most $0.499n$ different remainders when divided by n only for $n = 4$ and for odd primes n . Indeed, assume that this happens and let $n > 2$ be arbitrary. Assume that n is not a power of 2 (the argument is similar in the other case) and pick an odd prime divisor p of n . If $f(k) \equiv r \pmod{n}$ for some $k, r \in \{0, 1, \dots, n-1\}$, then $f(\bar{k}) \equiv r \pmod{p}$, where \bar{k} is the remainder of k when divided by p . We deduce that \bar{r} can take at most $0.499p$ values, which means that r can take at most $0.499p \cdot \frac{n}{p} = 0.499n$ values (since for any remainder $x \pmod{p}$ there are exactly $\frac{n}{p}$ numbers between 0 and $n-1$ that are congruent to $x \pmod{p}$).

We will prove now that

$$f(X) = 420(X^2 - 1)^2$$

is a solution of the problem. This clearly satisfies the desired condition for $n = 4$, so it remains to check it when $n = p$ is an odd prime. This is clear for $p < 11$, so assume that $p \geq 11$. It suffices to prove that $(x^2 - 1)^2$ gives at most $0.499p$ remainders mod p when x varies over all residues mod p . Note that all $(x^2 - 1)^2$ are quadratic residues, and if y^2 is a quadratic residue, then y^2 is not of the form $(x^2 - 1)^2$ when $y + 1$ and $1 - y$ are quadratic non-residues. Letting N be the number of $y \in \{0, 1, \dots, p-1\}$ such that $1 \pm y$ are quadratic non-residues, we deduce that the numbers $(x^2 - 1)^2$ give at most $\frac{p+1}{2} - \frac{N}{2}$ different remainders mod p .

We still need to estimate N . Note that

$$N = \frac{1}{4} \sum_{y=2}^{p-2} \left(1 - \left(\frac{1+y}{p} \right) \right) \cdot \left(1 - \left(\frac{1-y}{p} \right) \right)$$

since for $2 \leq y \leq p-2$ the number $\frac{1}{4} \left(1 - \left(\frac{1+y}{p}\right)\right) \cdot \left(1 - \left(\frac{1-y}{p}\right)\right)$ equals 1 when $1 \pm y$ are quadratic non-residues and 0 otherwise. A brutal expansion gives

$$N = \frac{1}{4} \left(p - 3 - \sum_{y=2}^{p-2} \left(\frac{1-y}{p}\right) - \sum_{y=2}^{p-2} \left(\frac{1+y}{p}\right) + \sum_{y=2}^{p-2} \left(\frac{1-y^2}{p}\right) \right).$$

Next, we easily check that

$$\sum_{y=2}^{p-2} \left(\frac{1-y}{p}\right) = -1 - \left(\frac{2}{p}\right) = \sum_{y=2}^{p-2} \left(\frac{1+y}{p}\right)$$

and using proposition 5.110 we obtain

$$\sum_{y=2}^{p-2} \left(\frac{1-y^2}{p}\right) = -1 + \left(\frac{-1}{p}\right) \sum_{y=0}^{p-1} \left(\frac{y^2-1}{p}\right) = -1 + (-1)^{\frac{p+1}{2}}.$$

We deduce that

$$N = \frac{1}{4} \left(p - 2 + 2 \left(\frac{2}{p}\right) + (-1)^{\frac{p+1}{2}} \right) \geq \frac{p-5}{4}.$$

To conclude it, remains to check that

$$\frac{p+1}{2} - \frac{p-5}{8} \leq 0.499p$$

for $p \geq 11$, which is immediate. \square

We are now able to prove the following beautiful result, which will play a key role in the next section.

Theorem 5.117. (*V. Lebesgue*) *Let $p > 2$ be a prime and let n be an odd integer. The number of solutions of the congruence*

$$x_1^2 + \dots + x_n^2 \equiv 1 \pmod{p}$$

is $p^{n-1} + ((-1)^{\frac{p-1}{2}} p)^{\frac{n-1}{2}}$.

Proof. If n is any positive integer and a is an integer, let $N(a, n)$ be the number of solutions of the congruence $x_1^2 + \dots + x_n^2 \equiv a \pmod{p}$. Writing the congruence as

$$x_1^2 + \dots + x_{n-2}^2 \equiv a - (x_{n-1}^2 + x_n^2) \pmod{p},$$

we see that

$$N(a, n) = \sum_{x_{n-1}, x_n \in \{0, 1, \dots, p-1\}} N(a - x_{n-1}^2 - x_n^2, n-2).$$

By proposition 5.112, when x_{n-1}, x_n run over $\{0, 1, \dots, p-1\}$ the numbers $a - x_{n-1}^2 - x_n^2$ take each value \pmod{p} different from a exactly $p + (-1)^{\frac{p-1}{2}}$ times and take the value $a \pmod{p}$ exactly $p + (p-1)(-1)^{\frac{p+1}{2}}$ times. We deduce that

$$\begin{aligned} N(a, n) &= (p + (-1)^{\frac{p+1}{2}}) \sum_{b \neq a} N(b, n-2) + (p + (p-1)(-1)^{\frac{p-1}{2}}) N(a, n-2) \\ &= (p + (-1)^{\frac{p+1}{2}}) \sum_{b=0}^{p-1} N(b, n-2) + p(-1)^{\frac{p-1}{2}} N(a, n-2). \end{aligned}$$

Clearly $\sum_{b=0}^{p-1} N(b, n-2)$ counts $(n-2)$ -tuples of elements of $\{0, 1, \dots, p-1\}$, thus

$$\sum_{b=0}^{p-1} N(b, n-2) = p^{n-2}.$$

We conclude that

$$N(a, n) = p^{n-2}(p + (-1)^{\frac{p+1}{2}}) + p(-1)^{\frac{p-1}{2}} N(a, n-2).$$

Taking $a = 1$ in this last relation, an immediate induction on n finishes the proof of the theorem. \square

We will explain now an alternative (and perhaps more conceptual) way of proving the previous theorem, which has the advantage of being rather general and which also involves a certain number of very beautiful ideas. Since the

discussion to follow is a bit technical, the reader may safely skip this for a first reading.

Let N be the number of solutions of the congruence

$$x_1^2 + \dots + x_n^2 \equiv 1 \pmod{p}$$

and let $z = e^{\frac{2i\pi}{p}}$. The key observation is that for any integer a we have

$$1_{a \equiv 0 \pmod{p}} = \frac{1}{p} \sum_{k=0}^{p-1} z^{ka},$$

where the left-hand side equals 1 when $a \equiv 0 \pmod{p}$ and 0 otherwise. To prove this identity, note that it is trivial when $p \mid a$ and in the other case the formula for the sum of a geometric progression gives

$$\sum_{k=0}^{p-1} z^{ka} = \frac{1 - z^{pa}}{1 - z^a} = 0,$$

since $z^a \neq 1$ and $z^{pa} = (z^p)^a = 1$.

It follows that

$$N = \sum_{0 \leq x_1, \dots, x_n \leq p-1} \frac{1}{p} \sum_{k=0}^{p-1} z^{k(x_1^2 + \dots + x_n^2 - 1)},$$

in other words (by interchanging the sums)

$$N = \frac{1}{p} \sum_{k=0}^{p-1} z^{-k} \sum_{0 \leq x_1, \dots, x_n \leq p-1} z^{kx_1^2 + \dots + kx_n^2} = \frac{1}{p} \sum_{k=0}^{p-1} z^{-k} \left(\sum_{x=0}^{p-1} z^{kx^2} \right)^n.$$

The term for $k = 0$ is easy to evaluate and equals p^n . The other terms lead naturally to

Definition 5.118. Let

$$G(k) = \sum_{x=0}^{p-1} z^{kx^2} = \sum_{x=0}^{p-1} e^{\frac{2i\pi kx^2}{p}}, \quad G = G(1) = \sum_{x=0}^{p-1} z^{x^2}$$

the quadratic Gauss sum associated to k .

It turns out that all sums $G(k)$ can be easily expressed in terms of G :

Proposition 5.119. *If p does not divide k , then*

$$G(k) = \left(\frac{k}{p}\right) G.$$

Proof. If $k \equiv u^2 \pmod{p}$ for some nonzero u , then the remainders of $kx^2 = (ux)^2$ when divided by p are a permutation of the remainders of x^2 when x varies. Thus $G(k) = G$ is clear in this case. If k is not a square mod p , note that when x varies the numbers kx^2 reduced mod p cover 0 and twice each quadratic non-residue mod p . Thus in this case

$$G(k) = 1 + 2 \sum_{\left(\frac{x}{p}\right)=-1} z^x$$

and since

$$G = 1 + 2 \sum_{\left(\frac{x}{p}\right)=1} z^x,$$

the relation $G(k) = -G$ is equivalent to

$$1 + \sum_{\left(\frac{x}{p}\right)=-1} z^x + \sum_{\left(\frac{x}{p}\right)=1} z^x = 0.$$

But this is clear since the left-hand side is just $\sum_{x=0}^{p-1} z^x = 0$. □

Remark 5.120. The proof also shows that we have

$$G = \sum_{x=1}^{p-1} \left(\frac{x}{p}\right) z^x.$$

The key identity satisfied by G is the following.

Theorem 5.121. (*Gauss*) *We have*

$$G^2 = p(-1)^{\frac{p-1}{2}}.$$

In particular $|G| = \sqrt{p}$.

Proof. Using the previous proposition, we obtain (brutally expanding $G(k)^2$)

$$(p-1)G^2 = \sum_{k=1}^{p-1} G(k)^2 = \sum_{k=1}^{p-1} \sum_{x,y=0}^{p-1} z^{k(x^2+y^2)} = \sum_{x,y=0}^{p-1} \sum_{k=1}^{p-1} z^{k(x^2+y^2)}.$$

For fixed x, y , the sum $\sum_{k=1}^{p-1} z^{k(x^2+y^2)}$ equals -1 when p does not divide $k(x^2 + y^2)$ (or equivalently $x^2 + y^2$) and equals $p-1$ when $p \mid x^2 + y^2$. If $p \equiv 3 \pmod{4}$, the congruence $x^2 + y^2 \equiv 0 \pmod{p}$ has only the trivial solution $(x, y) = (0, 0)$ and so we obtain

$$(p-1)G^2 = p-1 - (p^2-1) = -p(p-1),$$

thus $G^2 = -p$ as desired. If $p \equiv 1 \pmod{4}$ the congruence $x^2 + y^2 \equiv 0 \pmod{p}$ has $2p-1$ solutions by proposition 5.112, thus we obtain in this case

$$(p-1)G^2 = (2p-1)(p-1) - (p^2-2p+1) = p(p-1)$$

and finally $G^2 = p$, as needed. \square

Remark 5.122. 1) One can also argue more directly as follows: brutally expand

$$G^2 = \sum_{x,y=0}^{p-1} z^{x^2+y^2}.$$

Proposition 5.112 shows that when x, y run from 0 to $p-1$ the numbers $x^2 + y^2$ cover every nonzero residue mod p exactly $p - (-1)^{\frac{p-1}{2}}$ times and cover the zero residue mod p exactly $p + (p-1)(-1)^{\frac{p-1}{2}}$ times. We conclude that

$$G^2 = p + (p-1)(-1)^{\frac{p-1}{2}} + (p - (-1)^{\frac{p-1}{2}})(z + z^2 + \dots + z^{p-1})$$

and the result follows from the equality $z + z^2 + \dots + z^{p-1} = -1$.

2) It follows from the previous theorem that $G = \pm\sqrt{p}$ when $p \equiv 1 \pmod{4}$ and $G = \pm i\sqrt{p}$ when $p \equiv 3 \pmod{4}$. Finding the correct sign is a very difficult problem that took several years for Gauss to solve! More precisely, Gauss proved that

$$G = \sqrt{p} \quad \text{if } p \equiv 1 \pmod{4} \quad \text{and} \quad G = i\sqrt{p} \quad \text{if not.}$$

Let us come back to our counting problem and recall that N is the number of solutions of the congruence $x_1^2 + \dots + x_n^2 \equiv 1 \pmod{p}$, where n is odd. We have already seen that

$$N = p^{n-1} + \frac{1}{p} \sum_{k=1}^{p-1} z^{-k} G(k)^n,$$

thus using the previous results and the fact that n is odd we obtain

$$\begin{aligned} N &= p^{n-1} + \frac{1}{p} \sum_{k=1}^{p-1} z^{-k} \left(\frac{k}{p}\right)^n G^n = p^{n-1} + \frac{1}{p} \left(\sum_{k=1}^{p-1} z^{-k} \left(\frac{k}{p}\right)\right) G^n \\ &= p^{n-1} + \frac{1}{p} \overline{G} G^n = p^{n-1} + G^{n-1} = p^{n-1} + ((-1)^{\frac{p-1}{2}} p)^{\frac{n-1}{2}}. \end{aligned}$$

This gives a different proof of Lebesgue's theorem 5.117. To fully appreciate the power of this approach, we suggest the reader to find an explicit formula for the number of solutions of any congruence of the form

$$a_1 x_1^2 + \dots + a_n x_n^2 \equiv b \pmod{p},$$

where a_1, \dots, a_n are integers not divisible by p and b is an integer. The next example discusses a special case.

Example 5.123. (MOSP) Let p be an odd prime. Find the number of 6-tuples (a, b, c, d, e, f) of integers between 0 and $p-1$ such that

$$a^2 + b^2 + c^2 \equiv d^2 + e^2 + f^2 \pmod{p}.$$

Proof. Let z be a primitive root of order p of unity. Arguing as in the previous discussion, it follows that the desired number of 6-tuples is

$$\begin{aligned} S &= \frac{1}{p} \sum_{0 \leq a, b, c, d, e, f \leq p-1} \sum_{k=0}^{p-1} z^{k(a^2+b^2+c^2-d^2-e^2-f^2)} \\ &= \frac{1}{p} \sum_{k=0}^{p-1} \sum_{0 \leq a, b, c, d, e, f \leq p-1} z^{k(a^2+b^2+c^2-d^2-e^2-f^2)} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{p} \sum_{k=0}^{p-1} \left(\sum_{0 \leq a \leq p-1} z^{ka^2} \right)^3 \cdot \left(\sum_{0 \leq d \leq p-1} z^{-kd^2} \right)^3 \\
&= p^5 + \frac{1}{p} \sum_{k=1}^{p-1} G(k)^3 \cdot \overline{G(k)^3} = p^5 + \frac{1}{p} \sum_{k=1}^{p-1} |G(k)|^6 = p^5 + (p-1)p^2,
\end{aligned}$$

since $|G(k)| = \left| \left(\frac{k}{p} \right) G \right| = |G| = \sqrt{p}$ for k not divisible by p . Hence the result is $p^5 + (p-1)p^2$. \square

5.4.3 The quadratic reciprocity law

We are now ready to give a simple proof of one of the cornerstones of number theory, the celebrated quadratic reciprocity law. This theorem (conjectured by Euler), one of the most beautiful in number theory, has hundreds of different proofs. It is certainly the most important result concerning quadratic residues.

Theorem 5.124. (*Gauss' quadratic reciprocity law*) For all **odd** primes $p \neq q$ we have

$$\left(\frac{p}{q} \right) \cdot \left(\frac{q}{p} \right) = (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}}.$$

Proof. Let N be the number of solutions of the congruence $x_1^2 + \dots + x_q^2 \equiv 1 \pmod{p}$. By Lebesgue's theorem 5.117

$$\begin{aligned}
N &= p^{q-1} + ((-1)^{\frac{p-1}{2}} p)^{\frac{q-1}{2}} = p^{q-1} + (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}} p^{\frac{q-1}{2}} \\
&\equiv 1 + (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}} \cdot \left(\frac{p}{q} \right) \pmod{q}.
\end{aligned}$$

If we could prove that $N \equiv 1 + \left(\frac{q}{p} \right) \pmod{q}$, then we would deduce that

$$\left(\frac{q}{p} \right) \equiv (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}} \cdot \left(\frac{p}{q} \right) \pmod{q}.$$

But then the difference between the two sides is a number between -2 and 2 , which is also divisible by $q > 2$, therefore it must be 0 and the quadratic reciprocity law follows.

We will prove now that

$$N \equiv 1 + \left(\frac{q}{p}\right) \pmod{q},$$

finishing the proof. The argument is purely combinatorial and very simple. Note that if (x_1, \dots, x_q) is a solution of the congruence $x_1^2 + \dots + x_q^2 \equiv 1 \pmod{p}$, then so are (x_2, \dots, x_q, x_1) , $(x_3, \dots, x_q, x_1, x_2), \dots$ and so we can create groups of q solutions of this equation, obtained by permuting cyclically x_1, \dots, x_q . Note that since q is a prime, the only possibility for two solutions in a group to be equal is to have $x_1 = \dots = x_q$. Thus if M is the number of solutions of the congruence which moreover satisfy $x_1 = \dots = x_q$, then $N \equiv M \pmod{q}$. It is fairly easy to determine M : this is the number of solutions of the congruence $qx_1^2 \equiv 1 \pmod{p}$, or equivalently $(qx_1)^2 \equiv q \pmod{p}$. Hence $M = 1 + \left(\frac{q}{p}\right)$ and so $N \equiv 1 + \left(\frac{q}{p}\right) \pmod{q}$, as desired. \square

We end the theoretical part of this section with a beautiful proof of the following key result.

Theorem 5.125. *For all odd primes p we have*

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}.$$

In particular, 2 is a quadratic residue mod p if and only if $\frac{p^2-1}{8}$ is even, which happens if and only if $p \equiv 1 \pmod{8}$ or $p \equiv -1 \pmod{8}$.

Proof. Note that $\frac{p^2-1}{8}$ is even if and only if $p \equiv \pm 1 \pmod{8}$, thus it suffices to prove the second statement. The identity

$$\left(\frac{p-1}{2}\right)! = 2 \cdot 4 \cdot 6 \cdot \dots \cdot 1 \cdot 3 \cdot 5 \cdot \dots$$

combined with the congruences

$$2j+1 \equiv -(p - (2j+1)) = -2 \cdot \left(\frac{p-1}{2} - j\right) \pmod{p}$$

give

$$\left(\frac{p-1}{2}\right)! \equiv 2 \cdot 4 \cdot 6 \cdot \dots \cdot (-2) \left(\frac{p-1}{2}\right) \cdot (-2) \left(\frac{p-1}{2} - 1\right) \cdot \dots \pmod{p}.$$

Consider now the case $p = 8k + 1$ for some k , then the previous congruence becomes

$$\begin{aligned} (4k)! &= \left(\frac{p-1}{2}\right)! \equiv 2 \cdot 4 \cdot \dots \cdot (4k) \cdot (-2) \cdot (4k) \cdot (-2) \cdot (4k-1) \cdot \dots \cdot (-2) \cdot (2k+1) \\ &= 2^{2k} (2k)! (-2)^{2k} (2k+1) \dots (4k) = 2^{4k} \cdot (4k)! \pmod{p}, \end{aligned}$$

which yields $2^{\frac{p-1}{2}} \equiv 1 \pmod{p}$ and so $\left(\frac{2}{p}\right) = 1$ by Euler's criterion. Similarly, if $p = 8k + 3$ the congruence becomes

$$\begin{aligned} (4k+1)! &= 2 \cdot 4 \cdot \dots \cdot (4k) \cdot (-2) \cdot (4k+1) \cdot (-2) \cdot (4k) \cdot \dots \cdot (-2) (2k+1) \\ &= 2^{2k} \cdot (2k)! \cdot (-2)^{2k+1} \cdot (2k+1) \dots (4k+1) = -2^{4k+1} \cdot (4k+1)! \pmod{p}, \end{aligned}$$

yielding $2^{\frac{p-1}{2}} = 2^{4k+1} \equiv -1 \pmod{p}$.

We deal similarly with the cases $p = 8k + 5$ and $p = 8k + 7$. □

Example 5.126. (Vietnam TST 2004) Prove that $2^n + 1$ does not have prime divisors of the form $8k - 1$ for any $n \geq 1$.

Proof. Suppose that $p \equiv -1 \pmod{8}$ and $p \mid 2^n + 1$ for some $n \geq 1$. Since $p \equiv 3 \pmod{4}$, n is odd (since otherwise $2^n + 1$ is of the form $x^2 + 1$). Then $2^n \equiv -1 \pmod{p}$ yields $2^{n+1} \equiv -2 \pmod{p}$ and so $\left(\frac{-2}{p}\right) = 1$. This is impossible, since $\left(\frac{-1}{p}\right) = -1$ and $\left(\frac{2}{p}\right) = 1$. The result follows. □

Example 5.127. (Romania TST 2005) Let $p \equiv 7 \pmod{8}$ be a prime. Prove that for all $n \geq 1$ we have

$$\sum_{k=1}^{p-1} \left\{ \frac{k^{2^n}}{p} - \frac{1}{2} \right\} = \frac{p-1}{2},$$

where $\{x\} = x - \lfloor x \rfloor$ is the fractional part of the real number x .

Proof. Observe first that for any real number x we have

$$\left\{x - \frac{1}{2}\right\} = \frac{1}{2} + \{2x\} - \{x\}$$

since $\left\lfloor x - \frac{1}{2} \right\rfloor = \lfloor 2x \rfloor - \lfloor x \rfloor - 1$ (as the reader can easily check). Thus the problem is reduced to the identity

$$\sum_{k=1}^{p-1} \left\{ \frac{2k^{2^n}}{p} \right\} = \sum_{k=1}^{p-1} \left\{ \frac{k^{2^n}}{p} \right\}.$$

Recalling that $p \left\{ \frac{x}{p} \right\}$ is the remainder of x when divided by p (when x is an integer), we reduced the problem to a statement about the remainders of the numbers k^{2^n} and $2k^{2^n}$. If we prove that there is an integer x such that $2 \equiv x^{2^n} \pmod{p}$, then we are done, as then the remainders of $2k^{2^n}$ (when k varies from 1 to $p-1$) are a permutation of the remainders of the numbers k^{2^n} for $1 \leq k \leq p-1$. Next, note that if $p \mid k^{2^n} - l^{2^n}$ for some $1 \leq k, l \leq p-1$, then $p \mid k^2 - l^2$ since $p \mid k^{\gcd(2^n, p-1)} - l^{\gcd(2^n, p-1)} = k^2 - l^2$. It follows that the remainders of the numbers k^{2^n} (when k varies) are a permutation of the quadratic residues mod p . Thus it suffices to prove that 2 is a quadratic residue mod p , which follows from $p \equiv -1 \pmod{8}$. \square

Example 5.128. (Romanian Masters in Mathematics 2013) If a is a positive integer, define $x_1 = a$ and $x_{n+1} = 2x_n + 1$. Find the largest positive integer k for which there is a positive integer a such that the numbers $2^{x_1} - 1, 2^{x_2} - 1, \dots, 2^{x_k} - 1$ are all primes.

Proof. Note that $k \geq 2$ since for $a = 2$ the numbers $2^{x_1} - 1 = 3$ and $2^{x_2} - 1 = 31$ are both primes. We will prove now that $k \leq 2$, by showing that for any $a \geq 1$ at least one of the numbers $2^{x_1} - 1, 2^{x_2} - 1, 2^{x_3} - 1$ is composite. Assume that these three numbers are all primes. It follows that $x_1 = a$, $x_2 = 2a + 1$, $x_3 = 4a + 3$ are also prime numbers. The case $a = 2$ is easy to settle (as then $2^{x_3} - 1 = 2^{11} - 1 = 23 \cdot 89$), so assume that a is an odd prime. Then $4a + 3 \equiv -1 \pmod{8}$, thus 2 is a quadratic residue mod $4a + 3$ and so $4a + 3 \mid 2^{\frac{4a+3-1}{2}} - 1 = 2^{x_2} - 1$. Since $2^{x_2} - 1$ is a prime, it follows that

$2^{2a+1} - 1 = 4a + 3$. This can be rewritten as $2^{2a-1} = a + 1$, and is clearly impossible since $2^{2a-1} \geq 1 + 2a - 1 = 2a > a + 1$. Thus the result of the problem is 2. \square

Example 5.129. Find all primes p such that $p! + p$ is a perfect square.

Proof. Clearly 2 and 3 are solutions of the problem. We will prove that these are the only solutions. Clearly $p = 5$ is not a solution, so let $p > 5$ be such that $p! + p = x^2$. Clearly x is odd, so $x^2 \equiv 1 \pmod{8}$ and then (as $p \geq 5$) $p \equiv 1 \pmod{8}$. If q is an odd prime smaller than p , then $q \mid p!$ and so

$$\left(\frac{p}{q}\right) = \left(\frac{p+p!}{q}\right) = 1.$$

Using the quadratic reciprocity law, we deduce that

$$\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}} = 1,$$

the last equality being a consequence of the congruence $p \equiv 1 \pmod{4}$. Thus all odd primes less than p are quadratic residues mod p . Since $p \equiv 1 \pmod{8}$, 2 is also a quadratic residue mod p .

We conclude that all numbers are quadratic residues mod p , which is absurd. Thus no $p > 3$ is a solution of the problem. \square

Example 5.130. Find all integers x, n such that $x^3 + 2x + 1 = 2^n$.

Proof. Clearly $n \geq 0$. If $n = 0$ we obtain $x = 0$, which gives us the solution $(x, n) = (0, 0)$. Clearly $n = 1$ gives no solution and $n = 2$ gives the solution $(x, n) = (1, 2)$. Assume now that $n \geq 3$, thus $8 \mid x^3 + 2x + 1$. Clearly x is odd, thus $x^3 \equiv x \pmod{8}$ and then $x \equiv 5 \pmod{8}$. Next, note that $2^n - 1 = x(x^2 + 2)$ is divisible by 3, thus n must be even. Finally, write the equation as

$$(x+1)(x^2 - x + 3) = 2^n + 2,$$

which shows that for any prime divisor p of $x^2 - x + 3$ we have $\left(\frac{-2}{p}\right) = 1$ and then $p \equiv 1, 3 \pmod{8}$. We deduce that $x^2 - x + 3 \equiv 1, 3 \pmod{8}$, which contradicts the fact that $x^2 - x + 3 \equiv 25 - 5 + 3 \equiv -1 \pmod{8}$. Thus the only solutions are $(x, n) = (0, 0), (1, 2)$. \square

Example 5.131. Prove that if r is an odd number, then there are infinitely many primes $p \equiv r \pmod{8}$.

Proof. Let us start with the case $r = 1$ and consider prime factors p of $n^4 + 1$, with $p \neq 2$. Then $p \mid (n^2)^2 + 1$, thus $p \equiv 1 \pmod{4}$. If $p \equiv 5 \pmod{8}$, then Fermat's little theorem yields

$$-1 = (-1)^{\frac{p-1}{4}} \equiv (n^4)^{\frac{p-1}{4}} = n^{p-1} \equiv 1 \pmod{p},$$

a contradiction. Thus $p \equiv 1 \pmod{8}$ for any such prime and the result follows now from Schur's theorem 4.67, which guarantees the existence of infinitely many p that divide a number of the form $n^4 + 1$.

Assume next that $r = 3$ and let $p_1 = 2, p_2 = 3, \dots$ be the sequence of primes. Consider $N_n = (p_2 p_3 \dots p_n)^2 + 2$ with $n > 2$. Then $N_n \equiv 3 \pmod{8}$, thus N_n must have a prime factor p not of the form $8k \pm 1$ (otherwise N_n would be congruent to $\pm 1 \pmod{8}$). Since $p \mid N_n$, -2 is a quadratic residue mod p , which yields $p \equiv 3 \pmod{8}$ (since p is not $1 \pmod{8}$). Also $p \neq 3$ (since $N_n \equiv 2 \pmod{3}$) and $p > p_n$. Varying n yields the desired result.

Similarly, if $r = 5$ one considers the number $N_n = (p_2 \dots p_n)^2 + 4 \equiv 5 \pmod{8}$ and argues as above, while if $r = 7$ one considers $2(p_1 p_2 \dots p_n)^2 - 1$. \square

Example 5.132. (AMM E 3012) Let a and b be positive integers such that $a > 1$ and $a \equiv b \pmod{2}$. Prove that $2^a - 1$ is not a divisor of $3^b - 1$.

Proof. The result is clear if a is even (as then $3 \mid 2^a - 1$), so assume that a and b are odd. If p is any prime factor of $2^a - 1$, then $2^a \equiv 1 \pmod{p}$ yields $\left(\frac{2}{p}\right) = 1$ and $3^b \equiv 1 \pmod{p}$ yields $\left(\frac{3}{p}\right) = 1$. The first relation holds if and only if $p \equiv \pm 1 \pmod{8}$. The relation $\left(\frac{3}{p}\right) = 1$ is equivalent (by the quadratic reciprocity law) to $(-1)^{\frac{p-1}{2}} \cdot \left(\frac{p}{3}\right) = 1$. Discussing two cases according to whether $p \equiv 1 \pmod{4}$ or $p \equiv 3 \pmod{4}$, one easily checks that the equality $(-1)^{\frac{p-1}{2}} \cdot \left(\frac{p}{3}\right) = 1$ is equivalent to $p \equiv \pm 1 \pmod{12}$. We deduce that $p \equiv \pm 1 \pmod{24}$ for any prime factor p of $2^a - 1$ and so $2^a - 1 \equiv \pm 1 \pmod{24}$. Since this is obviously impossible, the result follows. \square

Remark 5.133. In particular $2^n - 1$ cannot divide $3^n - 1$ unless $n = 1$.

Example 5.134. (Bulgaria 1998) Suppose that m, n are positive integers such that $\frac{(m+3)^n+1}{3m}$ is an integer. Prove that this integer is odd.

Proof. Assume that this integer is even, so that $6m$ divides $(m+3)^n+1$. First, observe that m is even (otherwise $(m+3)^n+1$ is odd). But then 4 divides $6m$, so it divides $(m+3)^n+1$, forcing $m \equiv 0 \pmod{4}$. Repeating the argument, we have $8|6m|(m+3)^n+1$. If 8 divides m , we would have $8|3^n+1$, which is not possible for any n . Thus $m \equiv 4 \pmod{8}$ and since 8 divides $(m+3)^n+1$, it follows that n is odd. For $m=4$ we can easily check the result, so assume that $m > 4$. Then there exists a prime $p > 2$ dividing m (as we proved that $m \equiv 4 \pmod{8}$). Then p divides 3^n+1 , thus -3 is a quadratic residue mod p (since n is odd and $3^{n+1} \equiv -3 \pmod{p}$). Using the quadratic reciprocity law, this implies that p is a quadratic residue mod 3 and so $p \equiv 1 \pmod{3}$. Since this happens for any $p > 2$ dividing m , it follows that we can write $m = 4k$ with $k \equiv 1 \pmod{3}$ and k odd. But then $m \equiv 1 \pmod{3}$, which makes impossible the divisibility $3|(m+3)^n+1$. The result follows. \square

Example 5.135. (Komal) Prove that there are infinitely many composite numbers of the form $2^{2^n}+1$ or $6^{2^n}+1$.

Proof. We will prove that if $2^{2^n}+1$ is a prime $p > 5$ for some n , then necessarily $6^{\frac{p-1}{2}}+1$ (which is still of the form $6^{2^m}+1$) is composite, more precisely a multiple of p (it is clear that it cannot be p , since it is greater than p). This is of course sufficient to conclude. Suppose that $p = 2^{2^n}+1$ is a prime > 5 and let us prove that $p \mid 6^{\frac{p-1}{2}}+1$. This is equivalent to $\left(\frac{6}{p}\right) = -1$, i.e. $\left(\frac{2}{p}\right) \cdot \left(\frac{3}{p}\right) = -1$. But since $p \equiv 1 \pmod{8}$, we have $\left(\frac{2}{p}\right) = 1$ and (using the quadratic reciprocity law) $\left(\frac{3}{p}\right) = (-1)^{\frac{p-1}{2}} \left(\frac{p}{3}\right) = -1$, since $p \equiv 2 \pmod{3}$ and $p \equiv 1 \pmod{4}$. The result follows. \square

Example 5.136. (Taiwan 2000) Prove that if m, n are integers greater than 1 such that $\varphi(5^m-1) = 5^n-1$, then $\gcd(m, n) > 1$.

Proof. Assume that $\gcd(m, n) = 1$. Then $\gcd(5^m-1, 5^n-1) = 4$. Note that we cannot find an odd prime p such that p^2 divides 5^m-1 . Indeed, if this

happened we would get $p|\varphi(5^m - 1)$, so that $p|5^n - 1$ and $p|5^m - 1$. But then $p = 2$, a contradiction. Thus we can write

$$5^m - 1 = 2^a p_1 \dots p_k, \quad 5^n - 1 = 2^{a-1} (p_1 - 1) \dots (p_k - 1)$$

for some $a \geq 2$ and some distinct odd primes p_1, \dots, p_k . Note that $k \geq 1$, since otherwise $5^m - 1 = 2^a$, $5^n - 1 = 2^{a-1}$ and so $a - 1 = 2$, which doesn't yield any solution. Thus 2^a divides $5^m - 1$ and $5^n - 1$, yielding $a \leq 2$ and then $a = 2$. It follows that 8 does not divide $5^m - 1$, forcing m to be odd. Combined with the fact that p_i divides $5^m - 1$, this implies that 5 is a quadratic residue mod p_i and using the quadratic reciprocity law we deduce that p_i is a quadratic residue mod 5. But then $p_i \equiv \pm 1 \pmod{5}$. Since $p_i - 1$ divides $5^n - 1$, we cannot have $p_i \equiv 1 \pmod{5}$, thus all p_i are congruent to -1 modulo 5. But then the equation $5^n - 1 = 2(p_1 - 1) \dots (p_k - 1)$ implies that $-1 = 2(-2)^k \pmod{5}$, while the equation $5^m - 1 = 4p_1 \dots p_k$ gives $-1 = (-1)^{k+1} \pmod{5}$. It is immediate to see that we cannot simultaneously have these two equations, finishing the solution. \square

5.5 Congruences involving rational numbers and binomial coefficients

In this relatively technical section we discuss a few more delicate congruences related to binomial coefficients. The reader is invited to skip this section for a first reading and to consult the following beautiful articles for further information: A. Granville, "Binomial coefficients modulo prime powers" and R. Mestrovic, "Lucas' theorem: its generalizations, extensions and applications".

5.5.1 Binomial coefficients modulo primes: Lucas' theorem

In this section we will discuss several results concerning the arithmetic of the binomial coefficients, more precisely we will try to discuss the remainder of $\binom{n}{k}$ when divided by a prime p , and use this to establish several rather remarkable congruences. The letter p will always denote a prime in this section.

We have already seen when discussing Fermat's little theorem how useful the congruence $p \mid \binom{p}{k}$ (for $1 \leq k < p$) is. Before dealing with more technical things, we would like to emphasize the very useful congruence below.

Proposition 5.137. *For all primes p and all $0 \leq k \leq p-1$ we have*

$$\binom{p-1}{k} \equiv (-1)^k \pmod{p}.$$

Proof. This follows directly from

$$\begin{aligned} k! \binom{p-1}{k} &= (p-k)(p-k+1)\dots(p-1) \equiv (-k)(-k+1)\dots(-1) \\ &\equiv (-1)^k k! \pmod{p} \end{aligned}$$

and the fact that $\gcd(k!, p) = 1$. □

The next problem establishes the converse of the previous proposition.

Example 5.138. Let $n > 1$ be an integer. Prove that if

$$\binom{n-1}{k} \equiv (-1)^k \pmod{n}$$

for all $k \in \{0, 1, \dots, n-1\}$, then n is a prime.

Proof. Assuming that this is not the case, let p be the smallest prime factor of n and write $n = rp$ for some $r > 1$. Then by assumption $\binom{n-1}{p} \equiv (-1)^p \pmod{n}$, thus

$$\frac{(n-1)(n-2)\dots(n-p)}{p!} \equiv (-1)^p \pmod{n}$$

and so

$$(n-1)(n-2)\dots(n-p+1)(r-1) \equiv (p-1)!(-1)^p \pmod{n}.$$

However the left-hand side is congruent to $(-1)^{p-1}(p-1)!(r-1) \pmod{n}$ and since p is the smallest prime factor of n we have $\gcd(n, (p-1)!) = 1$. Thus the previous congruence is equivalent to $(-1)^{p-1}(r-1) \equiv (-1)^p \pmod{n}$, that is $r \equiv 0 \pmod{n}$. This is clearly absurd and so n is a prime. □

We will attack now the general problem of understanding the remainder of $\binom{n}{k}$ when divided by a prime p . The final answer will be relatively complicated, so let us start with some simple but nontrivial observations. Consider the Euclidean division

$$n = pn_1 + n_2, \quad k = pk_1 + k_2$$

of n , respectively k by p , thus $n_1, k_1 \geq 0$ and $0 \leq n_2, k_2 < p$ are integers. The binomial coefficient $\binom{n}{k}$ is the coefficient of X^k in the polynomial $(1 + X)^n$. Since $p \mid \binom{p}{k}$ for $1 \leq k \leq p - 1$, we have $(1 + X)^p \equiv 1 + X^p \pmod{p}$ and so

$$(1 + X)^n = [(1 + X)^p]^{n_1} \cdot (1 + X)^{n_2} \equiv (1 + X^p)^{n_1} \cdot (1 + X)^{n_2} \pmod{p}.$$

The coefficient of $X^k = X^{pk_1+k_2}$ in $(1 + X^p)^{n_1} \cdot (1 + X)^{n_2}$ is $\binom{n_1}{k_1} \cdot \binom{n_2}{k_2}$ (with the usual convention that $\binom{a}{b} = 0$ whenever $a < b$) since the only way to write $k = pk_1 + k_2$ in the form $pu + v$ with $0 \leq u \leq n_1$ and $0 \leq v \leq n_2$ is by setting $u = k_1$ and $v = k_2$, if possible (i.e. if $k_1 \leq n_1$ and $k_2 \leq n_2$). The previous polynomial congruence yields therefore the following very useful result below.

Theorem 5.139. *If $n = pn_1 + n_2$ and $k = pk_1 + k_2$ for some integers $n_1, k_1 \geq 0$ and $0 \leq n_2, k_2 < p$, then*

$$\binom{n}{k} \equiv \binom{n_1}{k_1} \cdot \binom{n_2}{k_2} \pmod{p}.$$

We can consider the previous theorem as a recursive recipe of computing the remainder of $\binom{n}{k}$ when divided by p . Iterating this result yields the following classical and important theorem of Lucas. Before stating it, we recall¹ that for any integer $a > 1$ one can write any integer $n \geq 1$ uniquely in the form

$$n = n_0 + n_1a + n_2a^2 + \dots + n_ka^k$$

with $n_0, \dots, n_k \in \{0, 1, \dots, a-1\}$ and $n_k \neq 0$. This is called the base a expansion of n (when $a = 10$ we obtain the usual decimal expansion of positive integers) and the numbers n_0, n_1, \dots, n_k are called the digits of n when written in base a (for instance n_0 is simply the remainder of n when divided by a). We can now state and prove Lucas' theorem (we recall that $\binom{a}{b} = 0$ if $a < b$).

¹The reader not aware of this result is invited to prove it using the Euclidean division.

Theorem 5.140. (Lucas) Let $n = n_0 + n_1p + \dots + n_dp^d$ be the base p expansion of a positive integer n , and let $k \in \{0, 1, \dots, n\}$. Write² $k = k_0 + k_1p + \dots + k_dp^d$ for some integers $0 \leq k_1, \dots, k_d \leq p - 1$. Then

$$\binom{n}{k} \equiv \binom{n_0}{k_0} \cdot \binom{n_1}{k_1} \cdot \dots \cdot \binom{n_d}{k_d} \pmod{p}.$$

Proof. Applying the previous theorem several times yields

$$\begin{aligned} \binom{n}{k} &\equiv \binom{n_0}{k_0} \cdot \binom{n_1 + n_2p + \dots + n_dp^{d-1}}{k_1 + k_2p + \dots + k_dp^{d-1}} \\ &\equiv \binom{n_0}{k_0} \cdot \binom{n_1}{k_1} \cdot \binom{n_2 + \dots + n_dp^{d-2}}{k_2 + \dots + k_dp^{d-2}} \equiv \dots \equiv \binom{n_0}{k_0} \cdot \binom{n_1}{k_1} \cdot \dots \cdot \binom{n_d}{k_d} \pmod{p}. \end{aligned}$$

The result follows. \square

We illustrate now the previous theorem with a few examples.

Example 5.141. Prove that if n is a positive integer and p is a prime, then

$$\binom{n}{p} \equiv \left\lfloor \frac{n}{p} \right\rfloor \pmod{p}.$$

Proof. Writing $n = n_0 + n_1p + \dots + n_dp^d$ in base p , Lucas' theorem gives

$$\binom{n}{p} \equiv \binom{n_0}{0} \cdot \binom{n_1}{1} \cdot \binom{n_2}{0} \cdot \dots \cdot \binom{n_d}{0} = n_1 \equiv \left\lfloor \frac{n}{p} \right\rfloor \pmod{p},$$

which finishes the proof. \square

Example 5.142. (Fine's theorem, 1947) Let n be a positive integer and let n_0, \dots, n_d be the digits of n when written in base p , where p is a prime. Prove that the number of binomial coefficients not divisible by p in the n th row of Pascal's triangle is $(1 + n_0)(1 + n_1)\dots(1 + n_d)$.

²In other words we consider the base p expansion of k and add some leading zeroes if needed, in order to obtain the same number of digits in base p as n .

Proof. We need to find the number of integers $k \in \{0, 1, \dots, n\}$ for which p does not divide $\binom{n}{k}$. Write $k = k_0 + k_1p + \dots + k_dp^d$ for some $0 \leq k_i \leq p-1$ (uniquely determined by k). Then by Lucas's theorem

$$\binom{n}{k} \equiv \prod_{i=0}^d \binom{n_i}{k_i} \pmod{p},$$

thus p does not divide $\binom{n}{k}$ if and only if p does not divide any of the numbers $\binom{n_i}{k_i}$. Since $0 \leq k_i, n_i < p$, this happens precisely when $k_i \leq n_i$ for all $0 \leq i \leq d$. Thus for each $0 \leq i \leq d$ we have exactly $n_i + 1$ possibilities for k_i and since k is uniquely determined by the d -tuple (k_0, k_1, \dots, k_d) , the result follows. \square

Remark 5.143. For $p = 2$ we recover Glaisher's classical theorem (obtained in 1899): the number of odd entries in the n th row of Pascal's triangle is 2^s , where s is the number of 1's in the binary (i.e. base 2) expansion of n .

Example 5.144. Let p be a prime and let n be an integer greater than 1.

a) Prove that all binomial coefficients $\binom{n}{1}, \dots, \binom{n}{n-1}$ are divisible by p if and only if n is a power of p .

b) Prove that none of the binomial coefficients $\binom{n}{1}, \dots, \binom{n}{n-1}$ is divisible by p if and only if $n = qp^d - 1$ for some $0 < q < p$ and some $d \geq 0$. In particular $\binom{n}{1}, \dots, \binom{n}{n-1}$ are all odd if and only if $n + 1$ is a power of 2.

Proof. a) If $n = p^d$ for some $d \geq 1$, then clearly for all $k = k_0 + pk_1 + \dots + p^dk_d \in \{1, 2, \dots, n-1\}$ we have by Lucas's theorem

$$\binom{n}{k} \equiv \binom{0}{k_0} \cdot \dots \cdot \binom{0}{k_{d-1}} \cdot \binom{1}{k_d} \equiv 0 \pmod{p},$$

since $k_d = 0$ and at least one of the numbers k_0, \dots, k_{d-1} is positive. Conversely, suppose that $\binom{n}{1}, \dots, \binom{n}{n-1}$ are divisible by p , then Fine's theorem above gives $(1 + n_0)(1 + n_1)\dots(1 + n_d) = 2$ where n_0, \dots, n_d are the digits of n in base p . This immediately yields $n_0 = \dots = n_{d-1} = 0$ and $n_d = 1$, thus $n = p^d$ and the result follows.

b) If $n = qp^d - 1$ for some $d, \geq 0$, $0 < q < p$, then the base p expansion of n is

$$n = (q-1)p^d + (p-1)p^{d-1} + \dots + (p-1),$$

and the result follows directly from Lucas' theorem. Conversely, suppose that none of $\binom{n}{1}, \dots, \binom{n}{n-1}$ is divisible by p and write $n = n_0 + pn_1 + \dots + p^d n_d$ in base p . If $n_j < p - 1$ for some $j \in \{1, 2, \dots, d\}$, then $\binom{n}{(n_j+1)p^j}$ is divisible by p thanks to Lucas' theorem, and $1 \leq (n_j + 1)p^j < n$, a contradiction. Thus $n_0 = \dots = n_{d-1} = p - 1$ and the result follows immediately. \square

Example 5.145. (Iran TST 2012) Find all integers $n > 1$ such that for all $0 \leq i, j \leq n$ the numbers $i + j$ and $\binom{n}{i} + \binom{n}{j}$ have the same parity.

Proof. The condition is equivalent to the fact that the numbers $\binom{n}{i} - i$ have the same parity for $0 \leq i \leq n$. By taking $i = 0$, we see that they must be odd, thus the condition is equivalent to $\binom{n}{i} \equiv i + 1 \pmod{2}$ for $0 \leq i \leq n$. For $0 \leq i \leq n - 1$ we then have

$$\binom{n+1}{i+1} = \binom{n}{i+1} + \binom{n}{i} \equiv 2i + 3 \equiv 1 \pmod{2},$$

thus the numbers $\binom{n+1}{1}, \dots, \binom{n+1}{n}$ are all odd. By the previous example we obtain that $n + 2$ is a power of 2, thus $n = 2^k - 2$ for some $k \geq 2$. Conversely, for such n Lucas' theorem easily yields $\binom{n}{i} \equiv i + 1 \pmod{2}$ for $0 \leq i \leq n$: writing $n = 2^{k-1} + 2^{k-2} + \dots + 2$ and $i = i_{k-1}2^{k-1} + \dots + i_0$ gives

$$\binom{n}{i} \equiv \binom{1}{i_{k-1}} \cdot \dots \cdot \binom{1}{i_1} \cdot \binom{0}{i_0} \pmod{2}$$

and it is a simple matter to check that the last expression has the same parity as $i_0 + 1$, i.e. as $i + 1$. \square

Example 5.146. Let p be a prime and let $n > 1$ be an integer. Prove that p does not divide $\binom{2n}{n}$ if and only if all digits of n when written in base p belong to $\{0, 1, \dots, \frac{p-1}{2}\}$.

Proof. Let $2n = a_0 + pa_1 + \dots + p^d a_d$ be the base p representation of $2n$ and let $n = b_0 + pb_1 + \dots + p^d b_d$ be the base p representation of n (possibly completed with some leading zeros). Lucas' theorem shows that p does not divide $\binom{2n}{n}$ if and only if $a_i \geq b_i$ for all $0 \leq i \leq d$. We need to prove that this

is equivalent to $\max_{0 \leq j \leq d} b_j \leq \frac{p-1}{2}$. Clearly this last condition is equivalent to $a_j = 2b_j$ for $0 \leq j \leq d$, so we obtain one implication. For the other implication, assume that $a_j \geq b_j$ for $0 \leq j \leq d$ and let us prove that $a_j = 2b_j$ for $0 \leq j \leq d$. Suppose that for some j we know that $a_j \equiv 2b_j \pmod{p}$, then $p > a_j - 2b_j \geq -b_j > -p$ and so necessarily $a_j = 2b_j$. On the other hand, we have

$$0 = (a_0 - 2b_0) + (a_1 - 2b_1)p + \dots + (a_d - 2b_d)p^d.$$

Thus $a_0 \equiv 2b_0 \pmod{p}$ and the previous discussion gives $a_0 = 2b_0$. Next, the previous relation yields $a_1 \equiv 2b_1 \pmod{p}$, thus $a_1 = 2b_1$. Continuing like this yields the desired result. \square

Example 5.147. (Vietnam TST 2010) Prove that $\binom{4n}{2n} + 1$ is not divisible by 3 for any positive integer n .

Proof. Assume that 3 divides $\binom{4n}{2n} + 1$ for some $n \geq 1$. Using the previous example, we deduce that in the base 3 representation $2n = a_0 + 3a_1 + \dots + 3^d a_d$ of $2n$ we have $a_i \in \{0, 1\}$ for all i , thus the base 3 representation of $4n$ is $(2a_0) + (2a_1) \cdot 3 + \dots + (2a_d) \cdot 3^d$. Lucas' theorem and the hypothesis then give

$$-1 \equiv \binom{4n}{2n} \equiv \prod_{j=0}^d \binom{2a_j}{a_j} \pmod{3}.$$

Note that $\binom{2a_j}{a_j}$ is congruent to -1 modulo 3 when $a_j = 1$ and to 1 otherwise. Thus the number of $j \in \{0, 1, \dots, d\}$ for which $a_j = 1$ must be odd. But this is clearly impossible, since $2n = a_0 + 3a_1 + \dots + 3^d a_d$ is even, thus $a_0 + \dots + a_d$ is even. The result follows. \square

5.5.2 Congruences involving rational numbers

By theorem 5.2 for any prime p and any $k \in \{1, 2, \dots, p-1\}$ the number $\frac{1}{p} \binom{p}{k}$ is an integer. A natural question is: what is the remainder mod p of this integer? In order to seriously study this question, we need to extend the notion of congruences from integers to certain rational numbers. Many of the more delicate results in the next section will crucially use such congruences.

We start by introducing a notion of congruence modulo p for rational numbers whose denominators are not multiples of p . This allows us to work with such fractions as with integers, which turns out to be extremely useful in practice. Let n be an integer greater than 1 and consider the subset of \mathbf{Q} defined by

$$\mathbf{Z}_{(n)} = \left\{ \frac{a}{b} \mid a, b \in \mathbf{Z}, \gcd(b, n) = 1 \right\}.$$

So $\mathbf{Z}_{(n)}$ consists of rational numbers whose denominator (when written in lowest terms) is relatively prime to n . Let us note that if $x, y \in \mathbf{Z}_{(n)}$ then xy , $x + y$ and $x - y$ are also in $\mathbf{Z}_{(n)}$, since if $x = \frac{a}{b}$ and $y = \frac{c}{d}$ then

$$xy = \frac{ac}{bd}, \quad x + y = \frac{ad + bc}{bd}, \quad x - y = \frac{ad - bc}{bd}$$

and $\gcd(bd, n) = 1$.

Definition 5.148. We say that $x, y \in \mathbf{Z}_{(n)}$ are congruent modulo n and write $x \equiv y \pmod{n}$ if $x - y = nz$ for some $z \in \mathbf{Z}_{(n)}$ or, equivalently, if the numerator of the fraction $x - y$ when written in lowest form is divisible by n .

The notion of congruence defined above extends the usual congruence on $\mathbf{Z} \subset \mathbf{Z}_{(n)}$ and has the same formal properties (see proposition 2.2), as the reader can easily check.

We make now the following important remark: if $x, y \in \mathbf{Z}$ then $x \equiv y \pmod{n}$ in $\mathbf{Z}_{(n)}$ is equivalent to $x \equiv y \pmod{n}$ in \mathbf{Z} . Indeed, the only non-trivial statement is that if $x \equiv y \pmod{n}$ in $\mathbf{Z}_{(n)}$, then $n \mid x - y$. But by assumption $x - y$ can be written as $\frac{na}{b}$ with $\gcd(a, b) = 1$ and $\gcd(n, b) = 1$. Since $x - y$ is an integer, it follows that $b \mid na$ and since $\gcd(b, na) = 1$, we obtain $b \mid 1$ and so $x - y = \pm na \in n\mathbf{Z}$.

Next, we make a very important observation concerning congruences with rational numbers, which turns out to be very handy in practice (as the next examples will illustrate). Let $x = \frac{a}{b} \in \mathbf{Z}_{(n)}$. By definition $\gcd(b, n) = 1$ and so there is a unique $c \in \{1, \dots, n-1\}$ such that $bc \equiv 1 \pmod{n}$. Then $x \equiv ac \pmod{n}$ in $\mathbf{Z}_{(n)}$. Indeed,

$$x - ac = \frac{a(1 - bc)}{b}$$

and the numerator is divisible by n , while the denominator is prime to n .

For instance, let us apply this observation to prove the following congruence (which will be improved in the next section to a congruence mod p^2 if $p > 3$)

$$1 + \frac{1}{2} + \dots + \frac{1}{p-1} \equiv 0 \pmod{p}$$

valid for any prime $p > 2$. Indeed, let $a_i \in \{1, 2, \dots, p-1\}$ be such that $ia_i \equiv 1 \pmod{p}$, then the previous discussion gives

$$1 + \frac{1}{2} + \dots + \frac{1}{p-1} \equiv a_1 + \dots + a_{p-1} \pmod{p}.$$

But since a_1, \dots, a_{p-1} are pairwise distinct modulo p , they are a permutation of $1, 2, \dots, p-1$ and so

$$a_1 + a_2 + \dots + a_{p-1} \equiv 1 + 2 + \dots + (p-1) = \frac{p(p-1)}{2} \equiv 0 \pmod{p}.$$

The same argument shows that for any prime p and for any positive integer k we have

$$1 + \frac{1}{2^k} + \dots + \frac{1}{(p-1)^k} \equiv 1 + 2^k + \dots + (p-1)^k \pmod{p}.$$

Using corollary 5.77 we obtain the beautiful and extremely useful congruence below.

Proposition 5.149. *For any prime p and any integer k which is not divisible by $p-1$ (in particular if $1 \leq k < p-1$) we have*

$$1 + \frac{1}{2^k} + \dots + \frac{1}{(p-1)^k} \equiv 0 \pmod{p}.$$

Before moving to concrete examples illustrating these relatively dry theoretical results, let us solve the original problem that motivated this short section: finding the remainder of $\frac{1}{p} \binom{p}{k}$ when divided by p .

Proposition 5.150. *For all primes p and all integers $1 \leq k \leq p-1$*

$$\frac{1}{p} \binom{p}{k} \equiv \frac{(-1)^{k-1}}{k} \pmod{p}.$$

Proof. This follows directly from the identity

$$\frac{1}{p} \binom{p}{k} = \frac{1}{k} \binom{p-1}{k-1}$$

and the congruence $\binom{p-1}{k-1} \equiv (-1)^{k-1} \pmod{p}$ (see proposition 5.137 for the latter). \square

It is now time to see how the previous results actually work in practice.

Example 5.151. Prove that for all primes $p > 3$

$$\sum_{j=1}^{\frac{p-1}{2}} \frac{1}{j^2} \equiv \sum_{j=0}^{\frac{p-3}{2}} \frac{1}{(2j+1)^2} \equiv 0 \pmod{p}.$$

Proof. These congruences follow directly from proposition 5.149 as follows:

$$0 \equiv \sum_{j=1}^{p-1} \frac{1}{j^2} = \sum_{j=1}^{\frac{p-1}{2}} \frac{1}{j^2} + \sum_{j=1}^{\frac{p-1}{2}} \frac{1}{(p-j)^2} \equiv 2 \sum_{j=1}^{\frac{p-1}{2}} \frac{1}{j^2} \pmod{p},$$

and

$$0 \equiv \sum_{j=1}^{p-1} \frac{1}{j^2} = \sum_{j=1}^{\frac{p-1}{2}} \frac{1}{(2j)^2} + \sum_{j=0}^{\frac{p-3}{2}} \frac{1}{(2j+1)^2} = \frac{1}{4} \sum_{j=1}^{\frac{p-1}{2}} \frac{1}{j^2} + \sum_{j=0}^{\frac{p-3}{2}} \frac{1}{(2j+1)^2}.$$

\square

Example 5.152. (Putnam 1996) Let p be a prime and let $k = \lfloor \frac{2p}{3} \rfloor$. Prove that

$$\binom{p}{1} + \binom{p}{2} + \dots + \binom{p}{k} \equiv 0 \pmod{p^2}.$$

Proof. Equivalently, we need to prove that

$$\sum_{j=1}^k \frac{1}{p} \binom{p}{j} \equiv 0 \pmod{p}.$$

But using proposition 5.150 we obtain

$$\sum_{j=1}^k \frac{1}{p} \binom{p}{j} \equiv \sum_{j=1}^k \frac{(-1)^{j-1}}{j} = \sum_{j=1}^k \frac{1}{j} - 2 \sum_{j=1}^{\lfloor \frac{k}{2} \rfloor} \frac{1}{2j} \equiv \sum_{j=1}^k \frac{1}{j} + \sum_{j=1}^{\lfloor \frac{k}{2} \rfloor} \frac{1}{p-j} \pmod{p}.$$

One easily checks that $p - \lfloor \frac{k}{2} \rfloor = k + 1$ by distinguishing the cases $p \equiv 1 \pmod{6}$ and $p \equiv 5 \pmod{6}$. Using proposition 5.149 we finally obtain

$$\sum_{j=1}^k \frac{1}{p} \binom{p}{j} \equiv \sum_{j=1}^{p-1} \frac{1}{j} \equiv 0 \pmod{p}. \quad \square$$

Example 5.153. Let p be an odd prime number. +Prove that

$$\sum_{i=1}^{p-1} \frac{2^i}{i} \equiv \sum_{i=1}^{\frac{p-1}{2}} \frac{1}{i} \equiv -\frac{2^p - 2}{p} \pmod{p}.$$

Proof. By proposition 5.150 we have

$$\sum_{i=1}^{p-1} \frac{2^i}{i} \equiv \sum_{i=1}^{p-1} 2^i \cdot \frac{(-1)^{i-1}}{p} \binom{p}{i} = \frac{2 - 2^p}{p} \pmod{p}.$$

On the other hand, let

$$A = \sum_{i=1}^{\frac{p-1}{2}} \frac{1}{i} \quad \text{and} \quad B = \sum_{i=1}^{\frac{p-1}{2}} \frac{1}{2i-1}.$$

We have

$$\frac{A}{2} + B = \sum_{i=1}^{p-1} \frac{1}{i} \equiv 0 \pmod{p},$$

hence

$$A \equiv \frac{A}{2} - B = \sum_{i=1}^{p-1} \frac{(-1)^i}{i} \pmod{p}.$$

Using again proposition 5.150 we obtain

$$A \equiv \sum_{i=1}^{p-1} \frac{(-1)^i \cdot (-1)^{i-1}}{p} \binom{p}{i} = \frac{2-2^p}{p} \pmod{p}$$

and the result follows. \square

Remark 5.154. A consequence of the proof is that for any odd prime p we have

$$\frac{2^{p-1}-1}{p} \equiv 1 + \frac{1}{3} + \dots + \frac{1}{p-2} \pmod{p}.$$

Example 5.155. (ELMO 2009) Let $p > 3$ be a prime and let x be an integer such that $p \mid x^3 - 1$ but $p \nmid x - 1$. Prove that

$$x - \frac{x^2}{2} + \frac{x^3}{3} - \dots - \frac{x^{p-1}}{p-1} \equiv 0 \pmod{p}.$$

Proof. By proposition 5.150 and the binomial formula we obtain

$$x - \frac{x^2}{2} + \frac{x^3}{3} - \dots - \frac{x^{p-1}}{p-1} \equiv \sum_{k=1}^{p-1} \frac{1}{p} \binom{p}{k} x^k = \frac{(1+x)^p - x^p - 1}{p} \pmod{p},$$

thus it suffices to prove that

$$(1+x)^p \equiv 1 + x^p \pmod{p^2}.$$

This follows from example 5.18, since by assumption $p \mid x^2 + x + 1$. \square

Example 5.156. (IMO Shortlist 2011) Let p be an odd integer. If $a \in \mathbf{Z}$, let

$$S_a = \frac{a}{1} + \frac{a^2}{2} + \dots + \frac{a^{p-1}}{p-1}.$$

Prove that if m, n are integers such that $S_3 + S_4 - 3S_2 = \frac{m}{n}$, then $p \mid m$.

Proof. Proposition 5.150 gives

$$\begin{aligned} S_a &= \sum_{k=1}^{p-1} \frac{a^k}{k} \equiv \frac{1}{p} \sum_{k=1}^{p-1} (-1)^{k-1} a^k \binom{p}{k} \\ &= -\frac{1}{p} \sum_{k=1}^{p-1} (-a)^k \binom{p}{k} = \frac{(a-1)^p - a^p + 1}{p} \pmod{p}, \end{aligned}$$

hence

$$\begin{aligned} S_3 + S_4 - 3S_2 &\equiv \frac{2^p - 3^p + 1 + 3^p - 4^p + 1 - 3 + 3 \cdot 2^p - 3}{p} \\ &= -\frac{(2^p - 2)^2}{p} \equiv 0 \pmod{p}, \end{aligned}$$

the last congruence being a consequence of Fermat's little theorem. \square

5.5.3 Higher congruences: Fleck, Morley, Wolstenholme,...

We will deal now with higher congruences (i.e. modulo powers of p) involving binomial coefficients. This will crucially use the previous two sections. The following beautiful and classical congruence due to Babbage (1819) is based on theorem 5.2 and the very important Vandermonde's identity

$$\binom{m+n}{k} = \sum_{i=0}^k \binom{m}{i} \cdot \binom{n}{k-i}, \quad (2)$$

which follows by identifying the coefficients of X^k in both sides of the equality

$$(1+X)^{m+n} = (1+X)^m \cdot (1+X)^n.$$

Example 5.157. Prove that for all primes p we have

$$\binom{2p}{p} \equiv 2 \pmod{p^2}.$$

Equivalently, $\binom{2p-1}{p-1} \equiv 1 \pmod{p^2}$ if $p > 2$ is a prime.

Proof. Vandermonde's identity specializes to

$$\binom{2p}{p} = \sum_{k=0}^p \binom{p}{k}^2.$$

Using theorem 5.2 we obtain $p^2 \mid \binom{p}{k}^2$ for $1 \leq k \leq p-1$, thus $\binom{2p}{p} \equiv 2 \pmod{p^2}$. The last assertion of the problem follows directly from what we have already done, since $\binom{2p-1}{p-1} = \frac{1}{2} \binom{2p}{p}$. \square

The next classical and important theorem improves the result established in the previous example and the $k=1$ case of proposition 5.149.

Theorem 5.158. (*Wolstenholme, 1862*) For all primes $p > 3$

$$\sum_{j=1}^{p-1} \frac{1}{j} \equiv 0 \pmod{p^2} \quad \text{and} \quad \binom{2p}{p} \equiv 2 \pmod{p^3}.$$

Proof. Note that

$$2 \sum_{j=1}^{p-1} \frac{1}{j} = \sum_{j=1}^{p-1} \left(\frac{1}{j} + \frac{1}{p-j} \right) = p \sum_{j=1}^{p-1} \frac{1}{j(p-j)}$$

and using proposition 5.149 we obtain

$$\sum_{j=1}^{p-1} \frac{1}{j(p-j)} \equiv \sum_{j=1}^{p-1} \frac{1}{-j^2} \equiv 0 \pmod{p},$$

whence the first part of the theorem.

For the second part, propositions 5.150 and 5.149 give

$$\frac{1}{p^2} \left(\binom{2p}{p} - 2 \right) = \sum_{k=1}^{p-1} \left(\frac{1}{p} \binom{p}{k} \right)^2 \equiv \sum_{k=1}^{p-1} \frac{1}{k^2} \equiv 0 \pmod{p}.$$

The result follows. \square

Remark 5.159. 1) Wolstenholme's theorem was generalized by Ljunggren (1949) to $\binom{pa}{pb} \equiv \binom{a}{b} \pmod{p^3}$ and by Jacobsthal (1952) to

$$\binom{pa}{pb} \equiv \binom{a}{b} \pmod{p^q}, \quad q = 3 + v_p(ab(a-b))$$

for $a > b > 0$ and $p > 3$. The proof of this last congruence is very difficult.

2) The congruence $\binom{2n}{n} \equiv 2 \pmod{n}$ can hold when n is composite and odd, for instance for $n = 29 \cdot 937$. Similarly the congruence $\binom{2n}{n} \equiv 2 \pmod{n^2}$ holds for $n = 16843^2$.

3) Primes p for which $\binom{2p}{p} \equiv 2 \pmod{p^4}$ are called Wolstenholme primes. The only such primes less than 10^9 are 16843 and 2124679. No prime p satisfying $\binom{2p}{p} \equiv 2 \pmod{p^5}$ is known (and probably there is no such prime).

Example 5.160. (APMO 2006) Let $p \geq 5$ be a prime and let r be the number of ways of placing p checkers on a $p \times p$ checkerboard so that not all checkers are in the same row (however they may all be in the same column). Prove that r is divisible by p^5 .

Proof. The problem is equivalent to the congruence

$$\binom{p^2}{p} - p \equiv 0 \pmod{p^5}$$

or, after dividing by p , to

$$\prod_{k=1}^{p-1} \left(\frac{p^2}{k} - 1 \right) \equiv 1 \pmod{p^4}.$$

A brutal expansion of the left-hand side shows that

$$\prod_{k=1}^{p-1} \left(\frac{p^2}{k} - 1 \right) \equiv (-1)^{p-1} + (-1)^{p-2} \sum_{k=1}^{p-1} \frac{p^2}{k} \pmod{p^4},$$

thus the problem is reduced to proving the congruence

$$\sum_{k=1}^{p-1} \frac{1}{k} \equiv 0 \pmod{p^2},$$

which follows from theorem 5.158. □

Remark 5.161. We leave it as a challenge for the reader to establish the congruence

$$\binom{p^3}{p^2} \equiv \binom{p^2}{p} \pmod{p^8}$$

for all primes $p \geq 5$.

Next, we will try to explain the proof of a beautiful but difficult congruence due to Morley. Example 5.153 can be seen as a way of computing the remainder of $2^{p-1} - 1$ modulo p^2 in terms of the harmonic numbers

$$H_n = 1 + \frac{1}{2} + \dots + \frac{1}{n}.$$

More precisely, the second congruence in that example says that if $p > 2$ is a prime, then

$$2^{p-1} \equiv 1 - \frac{p}{2} H_{\frac{p-1}{2}} \pmod{p^2}.$$

The next example pushes this further, to a congruence modulo p^3 . This is an intermediate (but interesting in its own right) step in the proof of Morley's congruence. It is much more challenging than the previous problem.

Example 5.162. Prove that if p is an odd prime, then

$$2^{p-1} \equiv 1 - \frac{p}{2} H_{\frac{p-1}{2}} + \frac{p^2}{8} H_{\frac{p-1}{2}}^2 \pmod{p^3}.$$

Proof. Recall the identity

$$(n+1)(n+2)\dots(n+n) = 2^n \cdot 1 \cdot 3 \cdot \dots \cdot (2n-1).$$

Choosing $n = \frac{p-1}{2}$ we obtain

$$\frac{1}{2^{\frac{p-1}{2}}} (p+1)(p+3)\dots(2p-2) = 2^{\frac{p-1}{2}} \cdot 1 \cdot 3 \cdot \dots \cdot (p-2),$$

that is

$$2^{p-1} = \frac{(p+1)(p+3)\dots(p+p-2)}{1 \cdot 3 \cdot \dots \cdot (p-2)} = \prod_{j=0}^{\frac{p-3}{2}} \left(1 + \frac{p}{2j+1}\right).$$

Expanding the right-hand side yields

$$2^{p-1} \equiv 1 + p \sum_{j=0}^{\frac{p-3}{2}} \frac{1}{2j+1} + p^2 \sum_{0 \leq j < k \leq \frac{p-3}{2}} \frac{1}{(2j+1)(2k+1)} \pmod{p^3}.$$

Now, by Wolstenholme's congruence (theorem 5.158)

$$p \sum_{j=0}^{\frac{p-3}{2}} \frac{1}{2j+1} \equiv -p \sum_{j=1}^{\frac{p-1}{2}} \frac{1}{2j} = -\frac{p}{2} H_{\frac{p-1}{2}} \pmod{p^3}.$$

It is thus sufficient to prove that

$$2 \sum_{0 \leq j < k \leq \frac{p-3}{2}} \frac{1}{(2j+1)(2k+1)} \equiv \frac{1}{4} H_{\frac{p-1}{2}}^2 \pmod{p}.$$

The left-hand side equals

$$\left(\sum_{j=0}^{\frac{p-3}{2}} \frac{1}{2j+1} \right)^2 - \sum_{j=0}^{\frac{p-3}{2}} \frac{1}{(2j+1)^2}$$

and using the congruences in theorem 5.158 and example 5.151 we see that this is indeed congruent to $\frac{1}{4} H_{\frac{p-1}{2}}^2$ modulo p . \square

We are now ready to establish the following beautiful result of Morley.

Theorem 5.163. (*Morley's congruence*) *If $p > 3$ is a prime, then*

$$(-1)^{\frac{p-1}{2}} \binom{p-1}{\frac{p-1}{2}} \equiv 4^{p-1} \pmod{p^3}.$$

Proof. Let $x = H_{\frac{p-1}{2}}$. A brutal expansion yields

$$(-1)^{\frac{p-1}{2}} \binom{p-1}{\frac{p-1}{2}} = \prod_{i=1}^{\frac{p-1}{2}} \frac{i-p}{i} = \prod_{i=1}^{\frac{p-1}{2}} \left(1 - \frac{p}{i} \right)$$

$$\equiv 1 - px + p^2 \sum_{1 \leq i < j \leq \frac{p-1}{2}} \frac{1}{ij} = 1 - px + \frac{p^2}{2} \left(x^2 - \sum_{j=1}^{\frac{p-1}{2}} \frac{1}{j^2} \right) \pmod{p^3}.$$

By example 5.151 we obtain

$$(-1)^{\frac{p-1}{2}} \binom{p-1}{\frac{p-1}{2}} \equiv 1 - px + \frac{p^2}{2} x^2 \pmod{p^3}.$$

On the other hand, by problem 5.162 we obtain

$$4^{p-1} \equiv (2^{p-1})^2 \equiv \left(1 - \frac{p}{2}x + \frac{p^2}{8}x^2\right)^2 \equiv 1 - px + \frac{p^2}{2}x^2 \pmod{p^3}$$

and the result follows. \square

We end this section with two challenging examples, which illustrate many of the ideas and techniques introduced in the previous sections.

Example 5.164. (Fleck's congruence, 1913) Let p be a prime, j an integer and $n \geq 1$. Prove that if $q = \left\lfloor \frac{n-1}{p-1} \right\rfloor$ then

$$\sum_{\substack{0 \leq m \leq n \\ p \mid m-j}} (-1)^m \binom{n}{m} \equiv 0 \pmod{p^q}.$$

Proof. We will prove the result by induction on q . If $q = 0$, there is nothing to prove, so assume that $q \geq 1$ and the result is known for $q-1$. In particular, the inductive hypothesis can be applied to $N = n - (p-1)$, since $\left\lfloor \frac{N}{p-1} \right\rfloor = q-1$. Thus we know that for any integer j we have

$$S_j := \sum_{\substack{0 \leq m \leq N \\ p \mid m-j}} (-1)^m \binom{N}{m} \equiv 0 \pmod{p^{q-1}}.$$

Using Vandermonde's identity and the congruence $\binom{p-1}{i} \equiv (-1)^i \pmod{p}$ (see proposition 5.137), we can then improve the previous congruences as follows

(for simplicity we no longer write the bounds on the indices, by using the convention that $\binom{a}{b} = 0$ whenever $b < 0$ or $a < b$)

$$\begin{aligned}
 \sum_{\substack{0 \leq m \leq n \\ p|m-j}} (-1)^m \binom{n}{m} &= \sum_{p|m-j} (-1)^m \binom{N+p-1}{m} \\
 &= \sum_{p|m-j} (-1)^m \sum_{i=0}^{p-1} \binom{p-1}{i} \binom{N}{m-i} = \sum_{i=0}^{p-1} (-1)^i \binom{p-1}{i} \sum_{p|m-j} (-1)^{m-i} \binom{N}{m-i} \\
 &= \sum_{i=0}^{p-1} (-1)^i \binom{p-1}{i} S_{j-i} \equiv \sum_{i=0}^{p-1} \sum_{p|m+i-j} (-1)^m \binom{N}{m} \pmod{p^q}.
 \end{aligned}$$

Note that the last sum is equal to $\sum_{m=0}^N (-1)^m \binom{N}{m} = 0$, hence the inductive step is proved and we are done. \square

Example 5.165. (Russia 2002) For each positive integer n , write

$$1 + \frac{1}{2} + \dots + \frac{1}{n} = \frac{A(n)}{B(n)},$$

where $A(n)$ and $B(n)$ are relatively prime integers. Prove that $A(n)$ is not a power of a prime for infinitely many n .

Proof. To simplify notations, write

$$f(n) = 1 + \frac{1}{2} + \dots + \frac{1}{n}.$$

Assume that there is N such that $A(n)$ is a power of a prime for all $n \geq N$. For each prime $p > N + 1$ we have $f(p-1) \equiv 0 \pmod{p^2}$ by Wolstenholme's theorem, thus $A(p-1)$ is a multiple of p^2 and must be a power of p , different from p .

This is the starting point of an induction that will show that $A(p^k - 1)$ is a power of p different from p for all $k \geq 1$. We have just proved this for $k = 1$, so assume that it holds for $k \geq 1$ and let us prove it for $k + 1$. We have

$$\frac{A(p^{k+1} - 1)}{B(p^{k+1} - 1)} = f(p^{k+1} - 1)$$

$$= \sum_{j=1}^{p^k-1} \frac{1}{pj} + \sum_{r=1}^{p-1} \sum_{j=0}^{p^k-1} \frac{1}{pj+r} = \frac{1}{p} f(p^k-1) + \sum_{r=1}^{p-1} \sum_{j=0}^{p^k-1} \frac{1}{pj+r}.$$

The term $\frac{1}{p} f(p^k-1)$ is $0 \pmod p$ by the inductive hypothesis. On the other hand, for all $1 \leq r \leq p-1$ we have

$$\sum_{j=0}^{p^k-1} \frac{1}{pj+r} \equiv \sum_{j=0}^{p^k-1} \frac{1}{r} \equiv 0 \pmod p.$$

We deduce that $A(p^{k+1}-1) \equiv 0 \pmod p$ and so $A(p^{k+1}-1)$ is a power of p . We still need to prove that $A(p^{k+1}-1)$ cannot be p . This will require the following nice observation: in general, if $2^j \leq n < 2^{j+1}$, then among the numbers $1, 2, \dots, n$ there is a unique multiple of 2^j (namely 2^j), thus 2^j divides $B(n)$ and so $B(n) > \frac{n}{2}$, yielding

$$A(n) > B(n) > \frac{n}{2}.$$

Therefore

$$A(p^{k+1}-1) > \frac{p^{k+1}-1}{2} \geq \frac{p^2-1}{2} > p,$$

which proves that $A(p^{k+1}-1)$ is not equal to p and finishes the induction.

We are now (finally!) almost done. Write $A(p^k-1) = p^{u_k}$ and note that since $A(p^k-1) > \frac{p^k-1}{2}$, we must have $u_k \geq k-1$, in particular the sequence $(u_k)_k$ tends to ∞ . On the other hand

$$\begin{aligned} f(p^k-1) &= 1 + \frac{1}{2} + \dots + \frac{1}{p^k-p} + \frac{1}{p^k-p+1} + \dots + \frac{1}{p^k-1} \\ &= f(p^k-p) + \frac{1}{p^k-p+1} + \dots + \frac{1}{p^k-1} \end{aligned}$$

and the sum in the right-hand side is $0 \pmod p$. We deduce that $A(p^k-p)$ is also a power of p , say $A(p^k-p) = p^{v_k}$. As above, the sequence $(v_k)_k$ tends to ∞ . It follows that

$$\frac{1}{p^k-p+1} + \dots + \frac{1}{p^k-1} = f(p^k-1) - f(p^k-p) \equiv 0 \pmod{p^{w_k}},$$

where $w_k = \min(u_k, v_k)$ tends to ∞ . Since

$$\frac{1}{p^k - p + 1} + \dots + \frac{1}{p^k - 1} \equiv - \left(1 + \frac{1}{2} + \dots + \frac{1}{p-1} \right) \pmod{p^k},$$

we deduce that for all k

$$1 + \frac{1}{2} + \dots + \frac{1}{p-1} \equiv 0 \pmod{p^{\min(w_k, k)}}.$$

This is certainly impossible, since $\min(w_k, k)$ tends to ∞ , while $1 + \frac{1}{2} + \dots + \frac{1}{p-1}$ is nonzero. \square

5.5.4 Hensel's lemma

In this section we study the congruence $f(x) \equiv 0 \pmod{p^n}$, where f is a polynomial with integer coefficients, p is a prime and $n > 1$ is an integer. Thanks to the previous sections, we already have a good understanding of congruences modulo primes, so it is natural to try to use this information in order to deal with congruences modulo higher powers of primes.

We argue inductively and assume that we already know how to solve the congruence $f(x) \equiv 0 \pmod{p^{n-1}}$. Let us fix a solution a of this latter congruence³ and try to understand the liftings of a to solutions of the congruence $f(x) \equiv 0 \pmod{p^n}$, i.e. those solutions y of this last congruence which also satisfy $y \equiv a \pmod{p^{n-1}}$. Write $y = a + p^{n-1}b$ for some integer b . Theorem 2.69 yields

$$f(y) = f(a + p^{n-1}b) \equiv f(a) + p^{n-1}bf'(a) \pmod{p^{2(n-1)}}$$

and since $2(n-1) \geq n$ we have $f(y) \equiv f(a) + p^{n-1}bf'(a) \pmod{p^n}$. Thus $y = a + p^{n-1}b$ is a solution of the congruence $f(x) \equiv 0 \pmod{p^n}$ if and only if

$$\frac{f(a)}{p^{n-1}} + bf'(a) \equiv 0 \pmod{p}.$$

³If there is no solution then clearly the congruence $f(x) \equiv 0 \pmod{p^n}$ also has no solution.

If $f'(a)$ is not divisible by p then there is a unique solution b of this linear congruence, hence there is a unique lifting of a to a solution of the congruence $f(x) \equiv 0 \pmod{p^n}$. Otherwise $p \mid f'(a)$ and we have two possibilities: either $p^n \mid f(a)$, in which case a lifts to p distinct solutions of the congruence $f(x) \equiv 0 \pmod{p^n}$ (namely all $a + p^{n-1}b$ with $0 \leq b \leq p-1$), or p^n does not divide $f(a)$, in which case a does not lift to any solution of $f(x) \equiv 0 \pmod{p^n}$. We summarize the previous discussion in the following important statement:

Theorem 5.166. (*Hensel's lemma*) *Let f be a polynomial with integer coefficients, p a prime and $n > 1$ an integer. Let a be a solution of the congruence $f(x) \equiv 0 \pmod{p^{n-1}}$. The number of solutions y of the congruence $f(x) \equiv 0 \pmod{p^n}$ satisfying $y \equiv a \pmod{p^{n-1}}$ is*

- 1 if p does not divide $f'(a)$.
- 0 if p divides $f'(a)$ and p^n does not divide $f(a)$.
- p if p divides $f'(a)$ and p^n divides $f(a)$.

The following consequence of the previous theorem appears quite often in practice.

Corollary 5.167. *Let f be a polynomial with integer coefficients and let p be a prime and $n > 1$ an integer. If $a \in \mathbf{Z}$ satisfies $f(a) \equiv 0 \pmod{p}$ and $\gcd(p, f'(a)) = 1$, then the congruence $f(x) \equiv 0 \pmod{p^n}$ has a unique solution b such that $b \equiv a \pmod{p}$.*

In other words the solution a of the congruence $f(x) \equiv 0 \pmod{p}$ lifts uniquely to a solution of the congruence $f(x) \equiv 0 \pmod{p^n}$, provided that p does not divide $f'(a)$.

Proof. Applying the previous theorem with $n = 2$ shows that a lifts uniquely to a solution a_1 of the congruence $f(x) \equiv 0 \pmod{p^2}$. Note that $f'(a_1) \equiv f'(a) \pmod{p}$, hence p does not divide $f'(a_1)$. Applying theorem 5.166 again shows that a_1 lifts uniquely to a solution a_2 of the congruence $f(x) \equiv 0 \pmod{p^3}$, and again p does not divide $f'(a_2)$. Repeating this process yields the desired result. \square

Example 5.168. Let p be an odd prime and let n be a positive integer.

- a) How many solutions does the congruence $x^{p-1} \equiv 1 \pmod{p^n}$ have?
- b) Answer the same question for the congruence $x^p \equiv 1 \pmod{p^n}$.

Proof. a) Consider the polynomial $f(X) = X^{p-1} - 1$. By Fermat's little theorem, the congruence $f(x) \equiv 0 \pmod{p}$ has $p-1$ solutions, namely $1, 2, \dots, p-1$. Moreover $f'(x)$ is relatively prime to p for any such x , thus by Hensel's lemma each solution of the congruence $f(x) \equiv 0 \pmod{p}$ uniquely lifts to one of the congruence $f(x) \equiv 0 \pmod{p^n}$. It follows that there are precisely $p-1$ solutions for all $n \geq 1$.

b) Letting $f(X) = X^p - 1$, the congruence $f(x) \equiv 0 \pmod{p}$ has one solution $x = 1$, again by Fermat's little theorem. This time we have $f'(1) \equiv 0 \pmod{p}$, so we cannot conclude easily as in part a). If $x^p \equiv 1 \pmod{p^n}$, then $x = 1 + py$ for some integer y , and the binomial formula allows us to rewrite the congruence in the form

$$y + \binom{p}{2}y^2 + \dots + p^{p-2}y^p \equiv 0 \pmod{p^{n-2}}.$$

If $n = 2$, this happens for all y , thus the congruence has p solutions in this case. Suppose that $n > 2$ and let $g(X) = X + \binom{p}{2}X^2 + \dots + p^{p-2}X^p$. Since $\binom{p}{2}, \dots, p^{p-2}$ are all multiples of p , the congruence $g(x) \equiv 0 \pmod{p}$ has only one solution $x = 0$ and $g'(0) = 1$ is not divisible by p . Hensel's lemma implies that $y = 0$ is the only solution of the congruence $g(y) \equiv 0 \pmod{p^{n-2}}$. Hence $x^p \equiv 1 \pmod{p^n}$ is equivalent to $x \equiv 1 \pmod{p^{n-1}}$, which shows that for all $n \geq 2$ the congruence $x^p \equiv 1 \pmod{p^n}$ has p solutions. \square

Remark 5.169. It would be much easier to deal with part b) using the lifting the exponent lemma: the congruence $x^p \equiv 1 \pmod{p^n}$ is equivalent to $v_p(x^p - 1) \geq n$, or (using that $x \equiv 1 \pmod{p}$ and the lifting the exponent lemma) $1 + v_p(x - 1) \geq n$, that is $p^{n-1} \mid x - 1$.

Let us see how the previous theoretical results work concretely in practice.

Example 5.170. Let p be a prime, a an integer relatively prime to p and n a positive integer. Consider the congruence $x^2 \equiv a \pmod{p^n}$.

a) Prove that if $p > 2$, then the congruence has exactly $1 + \left(\frac{a}{p}\right)$ solutions, i.e. two solutions when a is a quadratic residue modulo p and no solution otherwise.

b) Describe in terms of a and n the number of solutions of the congruence when $p = 2$.

Proof. Let $f(X) = X^2 - a$.

a) It is clear that if the congruence has solutions, then a must be a quadratic residue modulo p . Conversely, suppose that a is a quadratic residue modulo p . Then the congruence $f(x) \equiv 0 \pmod{p}$ has exactly two solutions and these solutions are relatively prime to p (recall that p does not divide a). Since p is odd, it follows that $\gcd(f'(x), p) = 1$ whenever $f(x) \equiv 0 \pmod{p}$. Hensel's lemma implies that the two solutions of the congruence $f(x) \equiv 0 \pmod{p}$ lift uniquely to solutions of the congruence $f(x) \equiv 0 \pmod{p^n}$, yielding the desired result.

b) It is clear that if $n = 1$ there is one solution, while if $n = 2$ there are no solutions unless $a \equiv 1 \pmod{4}$, in which case there are two solutions. Assume now that $n \geq 3$ and note that there is no solution unless $a \equiv 1 \pmod{8}$ (since $x^2 \equiv 1 \pmod{8}$ whenever x is odd). Thus assume that $a \equiv 1 \pmod{8}$ and let us prove first by induction that for all $k \geq 3$ the congruence $x^2 \equiv a \pmod{2^k}$ has solutions. This is clear for $k = 3$, so assume that $a \equiv x^2 \pmod{2^k}$ for an integer x . If $a \equiv x^2 \pmod{2^{k+1}}$ then we are done, otherwise $a \equiv x^2 + 2^k \pmod{2^{k+1}}$ and one easily checks that $a \equiv (x + 2^{k-1})^2 \pmod{2^{k+1}}$, yielding again the inductive step. Next, choose x_0 such that $x_0^2 \equiv a \pmod{2^n}$. Then $x^2 \equiv a \pmod{2^n}$ is equivalent to $x^2 \equiv x_0^2 \pmod{2^n}$ or $2^n \mid (x - x_0)(x + x_0)$. Since $\gcd(x - x_0, x + x_0) = 2$, this is also equivalent to $2^{n-1} \mid x - x_0$ or $2^{n-1} \mid x + x_0$, yielding four solutions in this case. \square

Example 5.171. Let p be an odd prime and let x be an integer relatively prime to p . Prove that $x^{\frac{p(p-1)}{2}} \equiv 1 \pmod{p^2}$ if and only if there is an integer y such that $y^2 \equiv x \pmod{p^2}$. How many integers $x \in \{0, 1, \dots, p^2 - 1\}$ have this property?

Proof. Suppose that $x^{\frac{p(p-1)}{2}} \equiv 1 \pmod{p^2}$, then $1 \equiv x^{\frac{p(p-1)}{2}} \equiv x^{\frac{p-1}{2}} \pmod{p}$, hence x is a quadratic residue modulo p . By example 5.170 there is an integer y such that $y^2 \equiv x \pmod{p^2}$, yielding one direction. Conversely, if such y exists then clearly x is a quadratic residue modulo p , hence $a := x^{\frac{p-1}{2}} \equiv 1 \pmod{p}$ and so

$$a^p = (1 + (a - 1))^p = 1 + p(a - 1) + \dots \equiv 1 \pmod{p^2},$$

yielding $x^{\frac{p(p-1)}{2}} \equiv 1 \pmod{p^2}$. It follows easily from Hensel's lemma (or even more directly from example 5.170) that the congruence $x^{\frac{p(p-1)}{2}} \equiv 1 \pmod{p^2}$ has $\frac{p(p-1)}{2}$ solutions (each solution modulo p lifts to p solutions modulo p^2). \square

Example 5.172. (ELMO Shortlist 2014) Is there an increasing infinite sequence of perfect squares $a_1 < a_2 < a_3 < \dots$ such that for all $k \geq 1$ we have that $13^k | a_k + 1$?

Proof. The answer is positive, and it suffices to prove that for each $k \geq 1$ the congruence $x^2 + 1 \equiv 0 \pmod{13^k}$ has solutions (as then there will be arbitrarily large values of x with $x^2 + 1 \equiv 0 \pmod{13^k}$, allowing an inductive construction of the desired sequence). Letting $f(x) = x^2 + 1$, the congruence $f(x) \equiv 0 \pmod{13}$ has a solution $x_0 = 5$ with $f'(x_0) = 10$ prime to 13, thus by Hensel's lemma this solution uniquely lifts to a solution of the congruence $f(x) \equiv 0 \pmod{13^k}$ for all k . The result follows. \square

Example 5.173. (IMO 1984) Find two positive integers a, b such that 7 does not divide $ab(a+b)$ but 7^7 divides $(a+b)^7 - a^7 - b^7$.

Proof. A first key point is factoring the expression $(a+b)^7 - a^7 - b^7$. For this it suffices to factor the polynomial $f(X) = (X+1)^7 - X^7 - 1$. Note that $f(0) = f(-1) = 0$, thus f is a multiple of $X(X+1)$. Also, if $z^3 = 1$ and $z \neq 1$ then $z+1 = -z^2$ and $f(z) = -z^{14} - z^7 - 1 = -z^2 - z - 1 = 0$. Thus f is also a multiple of $X^2 + X + 1$. Using this it is a simple matter to check that

$$f(X) = 7X(X+1)(X^2 + X + 1)^2.$$

Thus $7^7 | (a+b)^7 - a^7 - b^7$ if and only if $7^3 | a^2 + ab + b^2$ (using the fact that 7 does not divide $ab(a+b)$, by hypothesis). To make our life simpler we choose $a = 1$, so it suffices to find a positive integer b for which $7^3 | b^2 + b + 1$ (for any such b the number $b(b+1)$ is automatically not a multiple of 7). Letting $g(X) = X^2 + X + 1$ we need to study the congruence $g(x) \equiv 0 \pmod{7^3}$. We start by studying the congruence $g(x) \equiv 0 \pmod{7}$, which is easily seen to have two solutions, namely $x = 2$ and $x = 4$. Since $g'(2) = 5$ and $g'(4) = 9$ are nonzero modulo 7, we know by Hensel's lemma that each of these will lift to a unique solution modulo 7^3 , but since we are asked for a and b , we will

need to do the lifting. Let us lift the solution $x = 2$ to a solution modulo 7^2 . We are thus trying to find t such that $g(2 + 7t) \equiv 0 \pmod{7^2}$, or equivalently $g(2) + 7tg'(2) \equiv 0 \pmod{7^2}$. This is equivalent to $1 + 5t \equiv 0 \pmod{7}$ and the unique solution is $t = 4$, yielding a solution 30 of the congruence $g(x) \equiv 0 \pmod{7^2}$. Finally, we lift this solution to one modulo 7^3 , by looking for s such that $g(30 + 7^2s) \equiv 0 \pmod{7^3}$. This is equivalent to $g(30) + 7^2g'(30)s \equiv 0 \pmod{7^3}$, or $931 + 7^2 \cdot 61s \equiv 0 \pmod{7^3}$. This reduces to $19 + 61s \equiv 0 \pmod{7}$, or $5 - 2s \equiv 0 \pmod{7}$, with the unique solution $s = 6$. We obtain therefore the solution $30 + 7^2 \cdot 6 = 324$ of the congruence $g(x) \equiv 0 \pmod{7^3}$. Hence a solution of the problem is $a = 1$ and $b = 324$. Note that if we tried to lift the solution $x = 4$ of the congruence $g(x) \equiv 0 \pmod{7}$, we would have obtained the solution $b = 18$ of the congruence $g(x) \equiv 0 \pmod{7^2}$, which is also a solution of the congruence $g(x) \equiv 0 \pmod{7^3}$. \square

Example 5.174. (Putnam 2008) Let p be a prime and let $f \in \mathbb{Z}[X]$ be a polynomial. If $f(0), f(1), \dots, f(p^2 - 1)$ give distinct remainders when divided by p^2 , prove that $f(0), f(1), \dots, f(p^3 - 1)$ give distinct remainders when divided by p^3 .

Proof. Assume that $f(i) \equiv f(j) \pmod{p^3}$ for some i, j . Since $f(i) \equiv f(j) \pmod{p^2}$ and since f is injective mod p^2 , we deduce that $i \equiv j \pmod{p^2}$, say $j = i + p^2k$. It is enough to prove that $k \equiv 0 \pmod{p}$. Assume that this is not the case. We have

$$f(i) \equiv f(j) \equiv f(i + kp^2) \equiv f(i) + kp^2f'(i) \pmod{p^3},$$

so p divides $kf'(i)$, hence p divides $f'(i)$. But then

$$f(i + kp) \equiv f(i) + kpf'(i) \equiv f(i) \pmod{p^2},$$

which, combined with the hypothesis, yields $i + kp \equiv i \pmod{p^2}$, a contradiction. Thus $k \equiv 0 \pmod{p}$ and $i \equiv j \pmod{p^3}$. The result follows. \square

5.6 Problems for practice

Fermat's little theorem

1. Prove that for all primes p the number

$$\underbrace{11\dots1}_p \underbrace{22\dots2}_p \dots \underbrace{99\dots9}_p - \overline{12\dots9}$$

is divisible by p .

2. (Baltic Way 2009) Let p be a prime of the form $6k - 1$ and let a, b, c be integers such that $p \mid a + b + c$ and $p \mid a^4 + b^4 + c^4$. Prove that $p \mid a, b, c$.
3. (Poland 2010) Let p be an odd prime of the form $3k + 2$. Prove that

$$\prod_{k=1}^{p-1} (k^2 + k + 1) \equiv 3 \pmod{p}.$$

4. (Iran 2004) Let f be a polynomial with integer coefficients such that for all positive integers m, n there is an integer a such that $n \mid f(a^m)$. Prove that 0 or 1 is a root of f .
5. (Cippola, Rotkiewicz) Prove that if $n_1 > n_2 > \dots > n_k > 1$ are integers with $k > 1$ and $2^{n_k} > n_1$ then $F_{n_1} \dots F_{n_k}$ and $(2^{F_{n_1}} - 1) \dots (2^{F_{n_k}} - 1)$ are pseudo-primes, where $F_n = 2^{2^n} + 1$ is the n th Fermat number.
6. (India TST 2014) Find all polynomials f with integer coefficients such that $f(n)$ and $f(2^n)$ are relatively prime for all positive integers n .
7. (Rotkiewicz) An integer $n > 1$ is called pseudo-prime if n is composite and $n \mid 2^n - 2$. Prove that if p, q are distinct odd primes, then the following statements are equivalent:
 - a) pq is a pseudo-prime.
 - b) $p \mid 2^{q-1} - 1$ and $q \mid 2^{p-1} - 1$.
 - c) $(2^p - 1)(2^q - 1)$ is a pseudo-prime.

8. (Gazeta Matematica) Find all odd primes p for which $\frac{2^{p-1}-1}{p}$ is a perfect power.
9. (IMO Shortlist 2012) Define $\text{rad}(0) = \text{rad}(1) = 1$ and, for $n \geq 2$ let $\text{rad}(n)$ be the product of the different prime divisors of n . Find all polynomials $f(x)$ with nonnegative integer coefficients such that $\text{rad}(f(n))$ divides $\text{rad}(f(n^{\text{rad}(n)}))$ for all nonnegative integers n .
10. (Turkey TST 2013) Find all pairs of positive integers (m, n) such that

$$2^n + (n - \varphi(n) - 1)! = n^m + 1.$$

11. (Serbia 2015) Find all nonnegative integers x, y such that

$$(2^{2015} + 1)^x + 2^{2015} = 2^y + 1.$$

12. (Italy 2010) If n is a positive integer, let

$$a_n = 2^{n^3+1} - 3^{n^2+1} + 5^{n+1}.$$

Prove that infinitely many primes divide at least one of the numbers a_1, a_2, \dots

13. (China TST 2010) Find all positive integers $m, n \geq 2$, such that
- a) $m + 1$ is a prime number of the form $4k - 1$;
 - b) there is a prime number p and a nonnegative integer a such that

$$\frac{m^{2^n-1} - 1}{m - 1} = m^n + p^a.$$

Wilson's theorem

14. Let p be a prime. Prove that there is a positive integer n such that p is the smallest prime divisor of $n! + 1$.

15. Let $n > 1$ and suppose that there is $k \in \{0, 1, \dots, n-1\}$ such that

$$k!(n-k-1)! + (-1)^k \equiv 0 \pmod{n}.$$

Prove that n is a prime.

16. For each positive integer n find the greatest common divisor of $n! + 1$ and $(n+1)!$.
17. Let p be a prime and let a_1, a_2, \dots, a_{p-1} be consecutive integers.
- a) What are the possible remainders of $a_1 a_2 \dots a_{p-1}$ when divided by p ?
 - b) Suppose that $p \equiv 3 \pmod{4}$. Prove that a_1, \dots, a_{p-1} cannot be partitioned into two sets with the same product of their elements.
18. Find two primes p such that $(p-1)! + 1 \equiv 0 \pmod{p^2}$.
19. Find all sequences a_1, a_2, \dots of positive integers such that for all positive integers m, n

$$m! + n! \mid a_m! + a_n!.$$

20. Let p be an odd prime. A subset A of \mathbf{Z} is called a complete set of nonzero residue classes modulo p if A consists of $p-1$ integers giving pairwise distinct and nonzero remainders when divided by p . Prove that if $A = \{a_1, a_2, \dots, a_{p-1}\}$ and $B = \{b_1, b_2, \dots, b_{p-1}\}$ are complete sets of nonzero residue classes modulo p , then $\{a_1 b_1, \dots, a_{p-1} b_{p-1}\}$ is not a complete set of nonzero residue classes.
21. (Clement's criterion) Let n be an integer greater than 2. Prove that n and $n+2$ are both primes if and only if

$$4((n-1)! + 1) + n \equiv 0 \pmod{n(n+2)}.$$

22. Let $n > 1$ be an integer. Prove that there exists a positive integer k and $\varepsilon \in \{-1, 1\}$ such that $2k+1 \mid n + \varepsilon k!$.
23. (Moldova TST 2007) Prove that for infinitely many prime numbers p there is a positive integer n such that n does not divide $p-1$ and $p \mid n!+1$.

24. Find all polynomials f with integer coefficients such that for all primes p we have $f(p) \mid (p-1)! + 1$.
25. (adapted from Serbia 2010) Let a, n be positive integers such that $a > 1$ and $a^n + a^{n-1} + \dots + a + 1$ divides $a^{n!} + a^{(n-1)!} + \dots + a^{1!} + 1$. Prove that $n = 1$ or $n = 2$.

Lagrange's theorem and applications

26. Let p be a prime. Prove that the sequence of remainders mod p of the numbers $1, 2^2, 3^3, 4^4, \dots$ is periodic and find its least period.
27. (Don Zagier) Somebody incorrectly remembered Fermat's little theorem as saying that the congruence $a^{n+1} \equiv a \pmod{n}$ holds for all integers a . Describe the set of integers n for which this property is in fact true.
28. Let p be an odd prime. Find the largest degree of a polynomial f with the following properties:
- a) $\deg f < p$.
 - b) the coefficients of f are integers between 0 and $p-1$.
 - c) If m, n are integers and p does not divide $m-n$, then p does not divide $f(m) - f(n)$.
29. (Iran TST 2012) Let $p > 2$ be an odd prime. If $i \in \{0, 1, \dots, p-1\}$ and $f = a_0 + a_1X + \dots + a_nX^n$ is a polynomial with integer coefficients, we say that f is i -remainder if

$$\sum_{j>0, p-1 \nmid j} a_j \equiv i \pmod{p}.$$

Prove that the following statements are equivalent:

- a) f, f^2, \dots, f^{p-2} are 0-remainder and f^{p-1} is 1-remainder.
 - b) $f(0), f(1), \dots, f(p-1)$ form a complete residue system modulo p .
30. Find all integers $n > 2$ for which $n \mid 2^n + 3^n + \dots + (n-1)^n$.

31. (Alon, Dubiner) Let p be a prime and let $a_1, \dots, a_{3p}, b_1, \dots, b_{3p}$ be integers such that

$$\sum_{i=1}^{3p} a_i \equiv \sum_{i=1}^{3p} b_i \equiv 0 \pmod{p}.$$

Prove that there is a subset $I \subset \{1, 2, \dots, 3p\}$ with p elements such that

$$\sum_{i \in I} a_i \equiv \sum_{i \in I} b_i \equiv 0 \pmod{p}.$$

32. Prove that for any $n > 1$ the number $\binom{n}{0}^4 + \binom{n}{1}^4 + \dots + \binom{n}{n}^4$ is a multiple of any prime $p \in (n, \frac{4}{3}n]$.
33. Let f be a monic polynomial of degree $n \geq 1$, with integer coefficients. Suppose that b_1, \dots, b_n are pairwise distinct integers and that for infinitely many primes p the simultaneous congruences

$$f(x + b_1) \equiv f(x + b_2) \equiv \dots \equiv f(x + b_n) \equiv 0 \pmod{p}$$

have a common solution. Prove that the equations

$$f(x + b_1) = \dots = f(x + b_n) = 0$$

have a common integral solution.

34. (Romania TST 2016) Given a prime p , prove that

$$\sum_{k=1}^{\left\lfloor \frac{q}{p} \right\rfloor} k^{p-1}$$

is not divisible by q for all but finitely many primes q .

35. (China 2016) Let p be an odd prime and a_1, a_2, \dots, a_p be integers. Prove that the following two conditions are equivalent:

a) There is a polynomial P of degree $\leq \frac{p-1}{2}$ such that $P(i) \equiv a_i \pmod{p}$ for all $1 \leq i \leq p$;

b) For any $1 \leq d \leq \frac{p-1}{2}$

$$\sum_{i=1}^p (a_{i+d} - a_i)^2 \equiv 0 \pmod{p},$$

where indices are taken modulo p .

36. (USAMO 1999) Let p be an odd prime and let a, b, c, d be integers not divisible by p such that

$$\left\{ \frac{ra}{p} \right\} + \left\{ \frac{rb}{p} \right\} + \left\{ \frac{rc}{p} \right\} + \left\{ \frac{rd}{p} \right\} = 2$$

for all integers r not divisible by p (where $\{x\}$ is the fractional part of x). Prove that at least two of the numbers $a+b, a+c, a+d, b+c, b+d, c+d$ are divisible by p .

Quadratic residues and quadratic reciprocity

37. Let n be a positive integer such that $p = 4n + 1$ is a prime. Prove that $n^n \equiv 1 \pmod{p}$.
38. Let p be an odd prime. Prove that the number of integers $n \in \{1, 2, \dots, p-2\}$ such that n and $n+1$ are both quadratic residues mod p is $\frac{p-(-1)^{\frac{p-1}{2}}}{4} - 1$.
39. (Gazeta Matematică) Prove that for any $n \geq 1$ the number $3^n + 2$ does not have prime divisors of the form $24k + 13$.
40. Prove that there are infinitely many primes $p \equiv -1 \pmod{5}$.
41. Let $p = a^2 + b^2$ be an odd prime, with a, b positive integers and a odd. Prove that a is a quadratic residue mod p .
42. Let n be a positive integer and let a be a divisor of $36n^4 - 8n^2 + 1$, such that 5 does not divide a . Prove that the remainder of a when divided by 20 is 1 or 9.

43. Are there positive integers x, y, z such that $8xy = x + y + z^2$?

44. (Komal A 618) Prove that there are no integers x, y such that

$$x^3 - x + 9 = 5y^2.$$

45. Let p be an odd prime divisor of $n^4 - n^3 + 2n^2 + n + 1$, for some $n > 1$. Prove that $p \equiv 1, 4 \pmod{15}$.

46. Prove that infinitely many primes don't divide any of the numbers $2^{n^2+1} - 3^n$ with $n \geq 1$.

47. a) (Gauss) Prove that an odd prime p can be written $a^2 + 2b^2$ for some integers a, b if and only if $p \equiv 1, 3 \pmod{8}$.

b) (Euler, Lagrange) Prove that a prime $p \neq 3$ can be written $a^2 + 3b^2$ if and only if $p \equiv 1 \pmod{3}$.

48. (Moldova TST 2005) Let $f, g : \mathbf{N} \rightarrow \mathbf{N}$ be functions with the properties:

i) g is surjective;

ii) $2f(n)^2 = n^2 + g(n)^2$ for all positive integers n .

iii) $|f(n) - n| \leq 2004\sqrt{n}$ for all $n \in \mathbf{N}$.

Prove that f has infinitely many fixed points.

49. (Romania TST 2004) Let p be an odd prime and let

$$f(x) = \sum_{i=1}^{p-1} \left(\frac{i}{p}\right) X^{i-1}.$$

a) Prove that f is divisible by $X - 1$ but not by $(X - 1)^2$ if and only if $p \equiv 3 \pmod{4}$;

b) Prove that if $p \equiv 5 \pmod{8}$ then f is divisible by $(X - 1)^2$ but not by $(X - 1)^3$.

50. For an odd prime p , let $f(p)$ be the number of solutions of the congruence $y^2 \equiv x^3 - x \pmod{p}$.

a) Prove that $f(p) = p$ for $p \equiv 3 \pmod{4}$.

b) Prove that if $p \equiv 1 \pmod{4}$ then

$$f(p) \equiv (-1)^{\frac{p+3}{4}} \left(\frac{\frac{p-1}{2}}{\frac{p-1}{4}} \right) \pmod{p}.$$

c) For which primes p do we have $f(p) = p$?

51. Is there a polynomial f of degree 5 with integer coefficients such that f has no rational root and the congruence $f(x) \equiv 0 \pmod{p}$ has solutions for any prime p ?

52. Let p be an odd prime and let a be an integer not divisible by p . Let $N(a)$ be the number of solutions of the congruence $y^2 \equiv x^3 + ax \pmod{p}$ and let

$$S(a) = \sum_{k=0}^{p-1} \left(\frac{k^3 + ak}{p} \right).$$

1) Prove that $N(a) = p + S(a)$.

2) Prove that if $p \equiv 3 \pmod{4}$ then $S(a) = 0$ for all a , hence $N(a) = p$.

We assume from now on that $p \equiv 1 \pmod{4}$.

3) Prove that if b is not a multiple of p , then

$$S(ab^2) = \left(\frac{b}{p} \right) S(a).$$

4) Prove that

$$\sum_{a=0}^{p-1} S(a)^2 = 2p(p-1)$$

and that if $A = S(-1)$ and $B = S(a)$ for any quadratic non-residue a , then

$$A^2 + B^2 = 4p.$$

- 5) Prove that $A \equiv -(p+1) \pmod{8}$.
- 6) Deduce the following theorem of Jacobsthal: let $p \equiv 1 \pmod{4}$ be a prime and write $p = a^2 + b^2$ with a, b integers, a odd and $a \equiv -\frac{p+1}{2} \pmod{4}$. Then the congruence $y^2 \equiv x^3 - x \pmod{p}$ has $p+2a$ solutions.
53. (Mathematical Reflections) Find all primes p with the following property: whenever a, b, c are integers and $p \mid a^2b^2 + b^2c^2 + c^2a^2 + 1$, we also have $p \mid a^2b^2c^2(a^2 + b^2 + c^2 + a^2b^2c^2)$.

Congruences involving rational numbers and binomial coefficients

54. Let n be a positive integer and let $p \geq 2n+1$ be a prime. Prove that

$$\binom{2n}{n} \equiv (-4)^n \binom{\frac{p-1}{2}}{n} \pmod{p}.$$

55. (Mathematical Reflections O 96) Prove that if $q \geq p$ are primes, then

$$pq \mid \binom{p+q}{p} - \binom{q}{p} - 1.$$

56. (Hewgill) Let $n = n_0 + 2n_1 + \dots + 2^d n_d$ be the binary representation of an integer $n > 1$ and let S be the subset of $\{0, 1, \dots, n\}$ consisting of those k such that $\binom{n}{k}$ is odd. Prove that

$$\sum_{k \in S} 2^k = F_0^{n_0} F_1^{n_1} \dots F_d^{n_d},$$

where $F_k = 2^{2^k} + 1$ is the k th Fermat number.

57. (Calkin) Let a be a positive integer and let

$$x_n = \sum_{k=0}^n \binom{n}{k}^a$$

for $n \geq 1$. Let p be a prime, n an integer greater than 1 and let

$$n = n_0 + pn_1 + \dots + p^d n_d$$

be its base p representation. Prove that

$$x_n \equiv \prod_{i=0}^d x_{n_i} \pmod{p}.$$

58. Let p be a prime and let k be an odd integer such that $p - 1$ does not divide $k + 1$. Prove that

$$\sum_{j=1}^{p-1} \frac{1}{j^k} \equiv 0 \pmod{p^2}.$$

59. (Tuymaada 2012) Let $p = 4k + 3$ be a prime and write

$$\frac{1}{0^2 + 1} + \frac{1}{1^2 + 1} + \dots + \frac{1}{(p-1)^2 + 1} = \frac{m}{n}$$

for some relatively prime numbers m, n . Prove that $p \mid 2m - n$.

60. (IMO Shortlist 2012) Find all integers $m \geq 2$ such that $n \mid \binom{n}{m-2n}$ for any integer $n \in [\frac{m}{3}, \frac{m}{2}]$.

61. (Putnam 1991) Prove that for all odd primes p we have

$$\sum_{k=0}^p \binom{p}{k} \binom{p+k}{k} \equiv 2^p + 1 \pmod{p^2}.$$

62. (ELMO Shortlist 2011) Prove that if p is a prime greater than 3 then

$$\sum_{k=0}^{\frac{p-1}{2}} \binom{p}{k} 3^k \equiv 2^p - 1 \pmod{p^2}.$$

63. (IberoAmerican Olympiad 2005) Let $p > 3$ be a prime. Prove that

$$\sum_{i=1}^{p-1} \frac{1}{i^p} \equiv 0 \pmod{p^3}.$$

64. (AMM) Let $C_n = \frac{1}{n+1} \binom{2n}{n}$ be the n th Catalan number. Prove that

$$C_1 + C_2 + \dots + C_n \equiv 1 \pmod{3}$$

if and only if $n + 1$ has at least one digit equal to 2 in base 3.

65. Prove that for any prime $p > 5$ we have

$$\left(1 + p \sum_{k=1}^{p-1} \frac{1}{k}\right)^2 \equiv 1 - p^2 \sum_{k=1}^{p-1} \frac{1}{k^2} \pmod{p^5}.$$

66. (USA TST 2002) Let $p > 5$ be a prime number. For any integer x , define

$$f_p(x) = \sum_{k=1}^{p-1} \frac{1}{(px + k)^2}$$

Prove that $f_p(x) \equiv f_p(y) \pmod{p^3}$ for all positive integers x, y .

Chapter 6

p -adic valuations and the distribution of primes

The goal of this chapter is a rather detailed study of the p -adic valuation map $v_p : \mathbf{N} \rightarrow \mathbf{N}$ (where p is a fixed prime). Recall that if n is an integer greater than 1, then $v_p(n)$ is the exponent of p in the prime factorization of n . After reviewing the basic properties of the map v_p , we will use it to obtain results about the distribution of prime numbers.

6.1 The yoga of p -adic valuations

6.1.1 The local-global principle

Let us fix a prime number p . It will be convenient to extend the map $v_p : \mathbf{N} \rightarrow \mathbf{N}$ (whose definition was recalled above) to a map $v_p : \mathbf{Z} \rightarrow \mathbf{N} \cup \{\infty\}$ by setting $v_p(n) = v_p(|n|)$ for each $n \neq 0, \pm 1$, $v_p(\pm 1) = 0$ and $v_p(0) = \infty$. In other words, if n is a nonzero integer, then $v_p(n)$ is the largest nonnegative integer k such that p^k divides n . In particular $v_p(n) \geq 1$ is equivalent to $p \mid n$. We call $v_p(n)$ the p -adic valuation of n .

The following theorem summarizes the basic properties of the p -adic valuation map v_p . It is a direct consequence of the definition of this map and of the fundamental theorem of arithmetic.

Theorem 6.1. a) If n is a nonzero integer, then we can write $n = p^{v_p(n)} \cdot m$ with m relatively prime to p .

b) For each $n > 1$ we have

$$n = \prod_{p|n} p^{v_p(n)},$$

the product being taken over all primes p dividing n , or equivalently¹ over all prime numbers.

c) For all integers a, b we have

$$v_p(ab) = v_p(a) + v_p(b) \quad \text{and} \quad v_p(a+b) \geq \min(v_p(a), v_p(b)).$$

Proof. Parts a) and b) are clear from the fundamental theorem of arithmetic. Part c) is obvious if one of a, b is zero, so suppose that $ab \neq 0$. By a) we can write $a = p^{v_p(a)}u$ and $b = p^{v_p(b)}v$ with u, v relatively prime to p . Then uv is relatively prime to p and $ab = p^{v_p(a)+v_p(b)} \cdot (uv)$. Hence $v_p(ab) = v_p(a) + v_p(b)$. Next, $p^{\min(v_p(a), v_p(b))}$ divides both a and b , hence it divides $a+b$, hence $v_p(a+b) \geq \min(v_p(a), v_p(b))$. \square

The following crucial result shows that we can detect divisibility of integers by working "locally at every prime p ". This is the first local-global principle in number theory and we will use it a lot to prove divisibilities which would be rather difficult to prove otherwise.

Theorem 6.2. If a, b are integers then $a \mid b$ if and only if $v_p(a) \leq v_p(b)$ for all primes p .

Proof. We may assume that a, b are nonzero. If $a \mid b$ and $b = ac$ then $v_p(b) = v_p(a) + v_p(c) \geq v_p(a)$ for all p . Assume that $v_p(a) \leq v_p(b)$ for all p . Replacing a, b by their absolute values, we may assume that they are positive. Then $b = ac$, where $c = \prod_p p^{v_p(b)-v_p(a)}$, an integer. Hence $a \mid b$. \square

Remark 6.3. The previous theorem immediately implies the following result (which we have already proved using Gauss' lemma): if a, b are integers and

¹Since $p^{v_p(n)} = 1$ whenever p does not divide n .

$n \geq 1$ satisfies $a^n \mid b^n$, then $a \mid b$. Indeed, by the previous theorem we have for all primes p the inequality $nv_p(a) \leq nv_p(b)$. Thus $v_p(a) \leq v_p(b)$ for all p and the result follows by applying again the previous theorem.

We can also characterize n th powers of positive integers in terms of their p -adic valuations:

Theorem 6.4. *Let a and n be positive integers. Then a is the n th power of an integer if and only if $v_p(a) \equiv 0 \pmod{n}$ for all primes p (less formally, if and only if all exponents in the prime factorization of a are multiples of n).*

Proof. If $a = b^n$ is an n th power, then $v_p(a) = v_p(b^n) = nv_p(b) \equiv 0 \pmod{n}$ for all p . Conversely, if $v_p(a) = nb_p$ for all p and some nonnegative integers b_p , then $b_p = 0$ for all but finitely many primes p . If we set $b = \prod_p p^{b_p}$, then $b^n = \prod_p p^{nv_p(b)} = a$ and we are done. \square

Remark 6.5. This immediately implies the following result, which has already been proved using Gauss' lemma in a slightly tricky way: let a, b be relatively prime positive integers. If ab is the n th power of an integer, then a and b are n th powers of some integers. Indeed, suppose that $ab = c^n$ for some integer c . For all primes p we have $v_p(a) + v_p(b) = v_p(c^n) = nv_p(c) \equiv 0 \pmod{n}$. Moreover, since $\gcd(a, b) = 1$, p cannot divide both a and b , so we have $\min(v_p(a), v_p(b)) = 0$. We deduce that $v_p(a) \equiv v_p(b) \equiv 0 \pmod{n}$ for all primes p and the result follows from the previous theorem.

Finally, we compute the p -adic valuation of the greatest common divisor and least common multiple of two numbers (of course, they have obvious versions for several integers).

Proposition 6.6. *For all integers a, b we have*

$$v_p(\gcd(a, b)) = \min(v_p(a), v_p(b)) \quad \text{and} \quad v_p(\text{lcm}(a, b)) = \max(v_p(a), v_p(b)).$$

Proof. If one of the numbers a, b is 0 this is clear, so assume that $ab \neq 0$. Since $p^{\min(v_p(a), v_p(b))}$ divides both a and b , it divides $\gcd(a, b)$, hence

$$v_p(\gcd(a, b)) \geq \min(v_p(a), v_p(b)).$$

On the other hand, $p^{v_p(\gcd(a,b))}$ divides a and b , hence $v_p(\gcd(a,b)) \leq v_p(a)$ and $v_p(\gcd(a,b)) \leq v_p(b)$. The result follows. For lcm , use that $\text{lcm}(a,b) = \frac{ab}{\gcd(a,b)}$ to obtain

$$v_p(\text{lcm}(a,b)) = v_p(ab) - v_p(\gcd(a,b)) = v_p(a) + v_p(b) - \min(v_p(a), v_p(b)),$$

from which the result follows readily. \square

We end this section with a few concrete illustrations of the previous results.

Example 6.7. Prove that if $n > 1$ is an integer and p is a prime, then

$$v_p(\text{lcm}(1, 2, \dots, n)) = \lfloor \log_p(n) \rfloor.$$

Proof. The previous proposition gives

$$v_p(\text{lcm}(1, 2, \dots, n)) = \max_{1 \leq i \leq n} v_p(i).$$

Let $k = \lfloor \log_p(n) \rfloor$, so that $p^k \leq n < p^{k+1}$. Then clearly no $i \in \{1, 2, \dots, n\}$ is divisible by p^{k+1} and so

$$\max_{1 \leq i \leq n} v_p(i) = v_p(p^k) = k,$$

as desired. \square

Example 6.8. Prove that for all $n \geq 2$ we have

$$\text{lcm}(1, 2, \dots, n) \leq n^{\pi(n)},$$

where $\pi(n)$ is the number of primes not exceeding n .

Proof. If $p^k \leq n < p^{k+1}$, then $v_p(\text{lcm}(1, 2, \dots, n)) = k$ by example 6.7, hence

$$p^{v_p(\text{lcm}(1, 2, \dots, n))} \leq n.$$

The result follows by taking the product of these inequalities over all primes not exceeding n . \square

Example 6.9. Is there an infinite set of positive integers such that the sum of the elements in any nonempty subset is not a perfect power?

Proof. The answer is positive: consider the numbers $a_n = 2^n 3^{n+1}$ for $n \geq 1$ and let $A = \{a_1, a_2, \dots\}$. If $i_1 < i_2 < \dots < i_k$ are positive integers, then $x := a_{i_1} + a_{i_2} + \dots + a_{i_k}$ satisfies $v_2(x) = i_1$ and $v_3(x) = i_1 + 1$. Indeed, we have $x = 2^{i_1}y$ with $y = 3^{i_1+1} + 2^{i_2-i_1}3^{i_2+1} + \dots + 2^{i_k-i_1}3^{i_k+1}$ being odd, hence $v_2(x) = i_1$ and similarly $v_3(x) = i_1 + 1$. Since $\gcd(v_2(x), v_3(x)) = 1$, x cannot be a perfect power. Thus A has the desired property. \square

Example 6.10. (Saint Petersburg 2006) Let a_1, a_2, \dots, a_{101} be positive integers such that $\gcd(a_1, a_2, \dots, a_{101}) = 1$ and the product of any 51 of these numbers is divisible by the product of the remaining 50. Prove that $a_1 a_2 \dots a_{101}$ is a perfect square.

Proof. It suffices to prove that $v_p(a_1 \dots a_{101}) = \sum_{i=1}^{101} v_p(a_i)$ is even for all primes p . Fix a prime p and let $x_i = v_p(a_i)$. The hypothesis $\gcd(a_1, \dots, a_{101}) = 1$ yields $\min(x_1, \dots, x_{101}) = 0$. Assuming that $x_1 \geq x_2 \geq \dots \geq x_{101}$ (which we can do without loss of generality), we deduce that $x_{101} = 0$. Since $a_{51} a_{52} \dots a_{101}$ is a multiple of $a_1 \dots a_{50}$, we obtain

$$x_{51} + x_{52} + \dots + x_{100} + x_{101} \geq x_1 + x_2 + \dots + x_{50}.$$

However, $x_{101} = 0$ and $x_{51} \leq x_1, x_{52} \leq x_2, \dots, x_{100} \leq x_{50}$. Thus we must have $x_{51} = x_1, x_{52} = x_2, \dots, x_{100} = x_{50}$. We deduce that

$$x_1 + x_2 + \dots + x_{101} = 2(x_1 + \dots + x_{50})$$

is even, as desired. \square

Example 6.11. (Mathematical Reflections O 136) Let $(f_n)_{n \geq 1}$ be the Fibonacci sequence, i.e. $f_1 = f_2 = 1$ and $f_{n+1} = f_n + f_{n-1}$ for $n \geq 2$. Prove that $v_5(n) = v_5(f_n)$ for all n .

Proof. Let $x > y$ be the solutions of the equation $t^2 - t - 1 = 0$, so that

$$f_n = \frac{x^n - y^n}{\sqrt{5}}$$

for all $n \geq 1$. Then

$$f_{5n} = \frac{x^{5n} - y^{5n}}{\sqrt{5}} = f_n \cdot (x^{4n} + x^{3n}y^n + x^{2n}y^{2n} + x^ny^{3n} + y^{4n}).$$

Note that if we set

$$l_n = x^n + y^n,$$

the n th term of the Lucas sequence, then (using the fact that $xy = -1$)

$$\begin{aligned} x^{4n} + x^{3n}y^n + x^{2n}y^{2n} + x^ny^{3n} + y^{4n} &= x^{4n} + y^{4n} + (-1)^n(x^{2n} + y^{2n}) + 1 \\ &= (x^{2n} + y^{2n})^2 + (-1)^n(x^{2n} + y^{2n}) - 1 = l_{2n}^2 + (-1)^nl_{2n} - 1. \end{aligned}$$

Thus, setting

$$x_n = (-1)^nl_{2n} = (-x^2)^n + (-y^2)^n,$$

we have

$$f_{5n} = f_n \cdot (x_n^2 + x_n - 1).$$

We will now prove that $v_5(x_n^2 + x_n - 1) = 1$, which will yield $v_5(f_{5n}) = v_5(f_n) + 1$ and then $v_5(f_n) = v_5(n)$ by an immediate induction on $v_5(n)$ (using the fact that the sequence $(f_n)_{n \geq 1}$ is periodic modulo 5, with period 20, and that f_5, f_{10}, f_{15} are the only multiples of 5 among f_1, \dots, f_{19} , which can be easily checked by direct inspection). Note that it is enough to prove that $x_n \equiv 2 \pmod{5}$: if $x_n = 5k + 2$ then

$$x_n^2 + x_n - 1 = 25k^2 + 20k + 4 + 5k + 2 - 1 = 25(k^2 + k) + 5$$

and so clearly $v_5(x_n^2 + x_n - 1) = 1$. We will prove that $x_n \equiv 2 \pmod{5}$ by strong induction, the cases $n = 1$ and $n = 2$ being immediate. Next, note that $-x^2$ and $-y^2$ are solutions of the equation

$$(t + x^2)(t + y^2) = t^2 + 3t + 1 = 0$$

since $x^2y^2 = 1$ and $x^2 + y^2 = (x + y)^2 - 2xy = 3$. Thus the sequence $(x_n)_{n \geq 1}$ satisfies $x_{n+2} + 3x_{n+1} + x_n = 0$ for $n \geq 1$. In particular, if $x_n, x_{n+1} \equiv 2 \pmod{5}$, then $x_{n+2} \equiv -6 - 2 \equiv 2 \pmod{5}$. This finishes the proof.

Here is an alternate solution, suggested by Richard Stong. Let $l_0 = 2$, $l_1 = 1$, and $l_{n+1} = l_n + l_{n-1}$ for $n \geq 1$ be the Lucas sequence. Then from

$$\frac{l_0 + f_0\sqrt{5}}{2} = 1, \quad \frac{l_1 + f_1\sqrt{5}}{2} = \frac{1 + \sqrt{5}}{2} = \varphi,$$

and $\varphi^2 = \varphi + 1$, it follows by an easy induction that

$$\frac{l_n + f_n\sqrt{5}}{2} = \left(\frac{1 + \sqrt{5}}{2}\right)^n.$$

Hence by the binomial theorem (and the irrationality of $\sqrt{5}$),

$$2^{n-1}f_n = \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} \binom{n}{2k+1} 5^k = n + \sum_{k=1}^{\lfloor (n-1)/2 \rfloor} \frac{n}{2k+1} \binom{n-1}{2k} 5^k.$$

Since $5^k > 2k+1$, it follows that $v_5(2k+1) < k$, and hence every term in the sum is a multiple of $5^{v_5(n)+1}$. Thus we conclude that

$$v_5(f_n) = v_5(2^{n-1}f_n) = v_5(n). \quad \square$$

6.1.2 The strong triangle inequality

We have already established that if a, b are nonzero integers, then

$$v_p(a+b) \geq \min(v_p(a), v_p(b)),$$

in other words setting $|a|_p = p^{-v_p(a)}$ (we call $|a|_p$ the p -adic absolute value of a) we obtain

$$|a+b|_p \leq \max(|a|_p, |b|_p).$$

Note that this is much stronger than the usual triangle inequality

$$|a+b| \leq |a| + |b|$$

that holds for complex numbers a, b (with the usual absolute value). This is why the inequality

$$v_p(a+b) \geq \min(v_p(a), v_p(b))$$

is also sometimes called the strong triangle inequality.

The following theorem establishes a key property of the v_p map, related to the strong triangle inequality.

Theorem 6.12. *If p is a prime and a, b are integers such that $v_p(a) \neq v_p(b)$ then*

$$v_p(a + b) = \min(v_p(a), v_p(b)).$$

Proof. If $v_p(a) > v_p(b)$, then

$$a + b = p^{v_p(b)}(p^{v_p(a)-v_p(b)}u + v)$$

and p does not divide $p^{v_p(a)-v_p(b)}u + v$, since it does not divide v . Thus

$$v_p(a + b) = v_p(b) = \min(v_p(a), v_p(b)). \quad \square$$

We illustrate now these theoretical results with some rather interesting examples.

Example 6.13. (Czech-Slovak 2002) Let $m > 1$ be an integer. Prove that m is a perfect square if and only if for all positive integers n at least one of the numbers $(m + 1)^2 - m, (m + 2)^2 - m, \dots, (m + n)^2 - m$ is a multiple of n .

Proof. If $m = d^2$, then at least one of the numbers $m + 1 - d, m + 2 - d, \dots, m + n - d$ is a multiple of n , and the result follows. For the converse, pick a prime factor p of m and let $k = v_p(m)$. Choose $1 \leq i \leq p^{k+1}$ such that $p^{k+1} \mid (m + i)^2 - m$. If $v_p(m) \neq v_p((m + i)^2)$ then

$$k + 1 \leq v_p((m + i)^2 - m) = \min(v_p(m), v_p((m + i)^2)) \leq v_p(m) = k,$$

a contradiction. Hence $v_p(m) = v_p((m + i)^2) = 2v_p(m + i)$ is even and since this holds for any $p \mid m$, m is a perfect square. \square

Remark 6.14. The result still holds if we only assume that the statement of the problem holds for prime numbers n , but the proof is much more difficult.

We have already proved in theorem 4.67 that if f is a nonconstant polynomial with integer coefficients, then there are infinitely many primes p dividing a term of the sequence $f(1), f(2), \dots$. The following problem extends this result.

Example 6.15. (IMO Shortlist 2009) Let $f : \mathbf{N} \rightarrow \mathbf{N}$ be a nonconstant function such that $a - b$ divides $f(a) - f(b)$ for all $a, b \in \mathbf{N}$. Prove that there exist infinitely many primes p such that p divides $f(c)$ for some positive integer c .

Proof. Suppose that the conclusion fails and let p_1, \dots, p_k be all primes appearing in the prime factorizations of the numbers $f(1), f(2), \dots$. Take any positive integer x and write $f(x) = p_1^{\alpha_1} \dots p_k^{\alpha_k}$ for some nonnegative numbers $\alpha_1, \dots, \alpha_k$. Let $a_s = sp_1^{\alpha_1+1} \dots p_k^{\alpha_k+1}$ for $s \geq 1$. Since a_s divides $f(x + a_s) - f(x)$ and since $v_{p_i}(f(x)) < v_{p_i}(a_s)$, it follows that $v_{p_i}(f(x + a_s)) = v_{p_i}(f(x))$ for all i . But since all prime factors of $f(x + a_s)$ are among p_1, \dots, p_k , it follows that $f(x + a_s) = f(x)$, and this holds for all $s \geq 1$. But then $x + a_s - 1$ divides $f(x) - f(1) = f(x + a_s) - f(1)$ for all $s \geq 1$, so $f(x) = f(1)$. Since x was arbitrary, it follows that f is constant, contradicting the hypothesis of the problem. The result follows. \square

Example 6.16. (Kvant M 2163) Find all positive integers a and b such that :

- (i) $(a + b^2)(b + a^2)$ is a power of 2;
- (ii) $(a + b^3)(b + a^3)$ is a power of 3.

Proof. (i) We will prove that $a = b = 1$ is the unique solution of the problem. Assume that $(a, b) \neq (1, 1)$ and without loss of generality, that $a > 1$. Write $a + b^2 = 2^m$ and $b + a^2 = 2^n$ for some $m, n \geq 1$. If a is even, then so is b and since $v_2(a) < m = v_2(2^m)$ we have $v_2(2^m - a) = v_2(a)$, thus

$$2v_2(b) = v_2(b^2) = v_2(2^m - a) = v_2(a),$$

and similarly $2v_2(a) = v_2(b)$, contradicting our assumption that $v_2(a) > 0$.

Hence a is odd. If $b > 1$, then a similar argument as above yields

$$v_2(b + 1) < v_2(b^2 - 1) = v_2(2^m - (a + 1)) = v_2(a + 1)$$

and

$$v_2(a + 1) < v_2(a^2 - 1) = v_2(2^n - (b + 1)) = v_2(b + 1),$$

a contradiction. Hence $b = 1$ and $a + 1 = 2^m$, $a^2 + 1 = 2^n$. Since 4 does not divide $a^2 + 1$ for any integer a , we must have $n \leq 1$, contradiction with $a > 1$. Hence there are no solutions different from $a = b = 1$.

(ii) The solutions are $(a, b) = (1, 2)$ and $(a, b) = (2, 1)$. Assume that we have a solution with $a, b > 1$ and let $a^3 + b = 3^m$ and $a + b^3 = 3^n$. As above, if 3 divides a , then

$$3v_3(a) = v_2(3^m - b) = v_3(b)$$

and similarly $3v_3(b) = v_3(a)$, a contradiction with $v_3(a) > 0$. Hence $a \equiv 1, -1 \pmod{3}$. Note that if $a \equiv -1 \pmod{3}$, then $b \equiv 1 \pmod{3}$, thus by symmetry we may assume that $3 \mid a - 1$ and $3 \nmid b + 1$. Now if $a > 1$ a similar argument as above yields

$$v_3(a^3 - 1) = v_3(3^m - (b + 1)) = v_3(b + 1)$$

and

$$v_3(b^3 + 1) = v_3(3^n - (a - 1)) = v_3(a - 1).$$

Note that $v_3(a^3 - 1) > v_3(a - 1)$ and $v_3(b^3 + 1) > v_3(b + 1)$, since $a^2 + a + 1$ and $b^2 - b + 1$ are multiples of 3. Then the previous equalities yield $v_3(b + 1) > v_3(a - 1) > v_3(b + 1)$, a contradiction.

Hence we may assume that $a = 1$, so $b^3 + 1 = 3^n$ and $(b + 1)(b^2 - b + 1) = 3^n$. Suppose that $b > 2$, thus $n > 1$ and so $9 \mid b + 1$. Then $b^2 - b + 1 \equiv 3 \pmod{9}$ and since $b^2 - b + 1$ is a power of 3, we get $b^2 - b + 1 = 3$, a contradiction. Thus we must have $b = 2$ and the result follows. \square

The next two problems use a similar idea, which is a pretty subtle argument based on the pigeonhole principle and the strong triangle inequality.

Example 6.17. (IMO Shortlist 2011) Let d_1, d_2, \dots, d_9 be pairwise distinct integers. Prove that if x is a sufficiently large integer, then $(x + d_1)(x + d_2) \dots (x + d_9)$ has a prime divisor greater than 20.

Proof. Note that there are only 8 prime numbers less than 20, call them p_1, \dots, p_8 . By adding the same number to all d_i 's nothing is changed, so we may assume that $d_i > 0$ for all i . Now, assume that $(x + d_1) \dots (x + d_9)$ has all prime factors among p_1, \dots, p_8 , hence so do all numbers $x + d_1, \dots, x + d_9$. Assume that $x \geq (p_1 \dots p_8)^N$, with N sufficiently large. Then for each $1 \leq i \leq 9$ we can find $j_i \in \{1, 2, \dots, 8\}$ such that $v_{p_{j_i}}(x + d_i) \geq N$. Among the numbers $j_1, \dots, j_9 \in \{1, 2, \dots, 8\}$ two must be equal, say without loss of generality $j_1 = j_2$.

Then $p_{j_1}^N$ divides both $x + d_1$ and $x + d_2$, hence it divides $d_2 - d_1$. Since $d_2 \neq d_1$, this forces $p_{j_1}^N \leq |d_2 - d_1|$. Hence if N is chosen such that $2^N > \max_{i \neq j} |d_i - d_j|$, then for all $x > (p_1 \dots p_8)^N$ the number $(x + d_1) \dots (x + d_9)$ cannot have all of its prime factors among p_1, \dots, p_8 , and the problem is solved. \square

Example 6.18. (Erdős-Turan) Let $a_1 < a_2 < \dots$ be an infinite increasing sequence of positive integers. Prove that for any N we can find $i \neq j$ such that $a_i + a_j$ has a prime factor greater than N .

Proof. Fix N and let p_1, \dots, p_k be all primes not exceeding N . Suppose that for all $i \neq j$, all prime factors of $a_i + a_j$ are among p_1, \dots, p_k . Fix any positive integer d greater than all the numbers $a_v - a_u$ with $1 \leq u < v \leq k + 1$. Fix also $n > (p_1 \dots p_k)^d$ and note that for all $1 \leq i \leq k$ we have $a_n + a_i > (p_1 \dots p_k)^d$, thus there is $j_i \in \{1, 2, \dots, k\}$ such that $v_{p_{j_i}}(a_n + a_i) > d$. Since j_1, \dots, j_{k+1} are all between 1 and k , two of them must be equal, say $j_u = j_v$ with $1 \leq u < v \leq k + 1$. Let $p = p_{j_u}$, so that $v_p(a_n + a_u) > d$ and $v_p(a_n + a_v) > d$. It follows that $v_p(a_u - a_v) > d$, contradicting the fact that d is greater than $a_v - a_u$. \square

The next examples are more challenging.

Example 6.19. (Tuymaada 2004) Let a, n be positive integers such that $a \geq \text{lcm}(1, 2, \dots, n - 1)$. Prove that there are pairwise distinct prime numbers p_1, \dots, p_n such that $p_i \mid a + i$ for $1 \leq i \leq n$.

Proof. Let $b = \text{lcm}(1, 2, \dots, n - 1)$, thus $a \geq b$. Consider the numbers

$$x_i = \frac{a + i}{\gcd(a + i, b)}, \quad 1 \leq i \leq n.$$

We claim that x_1, \dots, x_n are pairwise relatively prime integers and $x_i > 1$ for all i . Note that this immediately implies the result, by taking p_i to be an arbitrary prime divisor of x_i . To prove the claim, note that $x_i > 1$ is clear, since the equality $a + i = \gcd(a + i, b)$ would force $a + 1 \leq b$. Assume now that a prime p divides both x_i and x_j , for some $1 \leq i < j \leq n$. Let $k = v_p(b)$. Then

$$\min(v_p(a + i), v_p(a + j)) \leq v_p((a + j) - (a + i)) = v_p(j - i) \leq v_p(b) = k.$$

We may assume that $v_p(a + i) \leq k$, but then

$$v_p(x_i) = v_p(a + i) - \min(v_p(a + i), k) = 0,$$

contradicting the fact that $p \mid x_i$. The result follows. \square

Example 6.20. (Iran TST 2013) Find all arithmetic progressions a_1, a_2, \dots of positive integers for which there is an integer $N > 1$ such that for all $k \geq 1$

$$a_1 a_2 \dots a_k \mid a_{N+1} a_{N+2} \dots a_{N+k}.$$

Proof. Write $a_n = a + nd$ for $n \geq 1$ and some $d \geq 1$. Note that if $a = 0$, then the sequence $(a_n)_n$ is a solution of the problem, since the product of k consecutive integers is a multiple of $k!$. We will prove that the case $a > 0$ is impossible. Dividing a and d by their greatest common divisor, we may assume that $\gcd(a, d) = 1$. For $k > N$ the divisibility condition can be rewritten as

$$a_1 a_2 \dots a_N \mid a_{k+1} a_{k+2} \dots a_{k+N},$$

by dividing the given divisibility relation by $a_{N+1} \dots a_k$. Note that $a_1 a_2 \dots a_N > N!$, hence there is a prime p such that $v_p(a_1 \dots a_N) > v_p(N!)$. Then p divides at least one of the numbers a_1, \dots, a_N , and these are all relatively prime to d since $\gcd(a, d) = 1$. Thus p does not divide d and so there is an integer $k > N$ such that $p^{v_p(a_1 \dots a_N)} \mid a_k = a + dk$. But then $v_p(a_k) > v_p(N!) \geq v_p(jd)$ for $1 \leq j \leq N$, hence

$$\begin{aligned} v_p(N!) &< v_p(a_1 \dots a_N) \leq v_p(a_{k+1} \dots a_{k+N}) \\ &= v_p((a_k + d)(a_k + 2d) \dots (a_k + Nd)) \\ &= v_p(a_k + d) + v_p(a_k + 2d) + \dots + v_p(a_k + Nd) \\ &= v_p(d) + v_p(2d) + \dots + v_p(Nd) = v_p(N!), \end{aligned}$$

a contradiction. \square

Example 6.21. (IMO 2010) Find all sequences of positive integers $(a_n)_{n \geq 1}$ such that $(a_n + m)(a_m + n)$ is a perfect square for all positive integers n, m .

Proof. It is clear that $a_n = n + k$ is a solution of the problem for all $k \geq 0$. We will prove that these are all solutions.

Let n, m be distinct positive integers and suppose that a prime p divides $a_n - a_m$. We will prove that $p \mid n - m$. We claim that we can find $s \geq 1$ such that $v_p(s + a_n)$ and $v_p(s + a_m)$ are odd. If the claim is proved, then $v_p(n + a_s)$ and $v_p(m + a_s)$ must be odd, since $(s + a_n)(n + a_s)$ and $(s + a_m)(m + a_s)$ are perfect squares. Thus p divides $n + a_s$ and $m + a_s$, and then $p \mid m - n$, as desired. Now, let us prove the existence of s . If $v_p(a_n - a_m) = 1$, choose $s = p^3 r - a_n$, where r is large enough and relatively prime to p . If $v_p(a_n - a_m) \geq 2$, choose $s = pr - a_n$, where r is large enough and prime to p .

Now, the previous paragraph shows that $a_n \neq a_m$ for all $n \neq m$, and also that $|a_n - a_{n+1}| = 1$. Thus $a_{n+1} - a_n$ and $a_{n+1} - a_{n+2}$ are both 1 or -1 , and distinct, thus they must add up to 0. This implies that $a_{n+2} - a_{n+1} = a_{n+1} - a_n$ for all $n \geq 1$, and since $a_n \geq 1$ for all n , we must have $a_{n+1} - a_n = 1$ for all n . Thus $a_n = n + k$ for some constant $k \geq 0$, and the problem is solved. \square

6.1.3 Lifting the exponent lemma

Let us start with some easy observations, which are however very useful in practice. Let a, b be integers and let p be a prime dividing $a - b$. Note that

$$a^p = (a - b + b)^p = (a - b)^p + p(a - b)^{p-1}b + \dots + p(a - b)b^{p-1} + b^p.$$

In the previous sum all terms except for the last one are multiples of p^2 , since $p \mid a - b$. We conclude that $p^2 \mid a^p - b^p$. In other words, if a and b are congruent mod p , then a^p and b^p are congruent mod p^2 , i.e. raising to p th power improves congruences! The same formula shows more generally that if p^l divides $a - b$ for some $l \geq 1$, then p^{l+1} divides $a^p - b^p$. This easily yields the following estimate.

Theorem 6.22. *Let a, b be integers and let p be a prime dividing $a - b$. Then for all positive integers c we have*

$$v_p(a^c - b^c) \geq v_p(a - b) + v_p(c), \quad \text{i.e.} \quad v_p\left(\frac{a^c - b^c}{a - b}\right) \geq v_p(c).$$

Proof. Let $k = v_p(c)$ and $l = v_p(a - b)$. Since $p^l \mid a - b$, the previous discussion shows that $p^{l+1} \mid a^p - b^p$, then $p^{l+2} \mid a^{p^2} - b^{p^2}$ and continuing like this we obtain $p^{l+k} \mid a^{p^k} - b^{p^k}$. Since $p^k \mid c$, we have $a^{p^k} - b^{p^k} \mid a^c - b^c$. Thus

$$v_p(a^c - b^c) \geq l + k = v_p(a - b) + v_p(c). \quad \square$$

Example 6.23. (Romania TST 2009) Let $a, n \geq 2$ be integers such that n divides $(a - 1)^k$ for some $k \geq 1$. Prove that n divides $1 + a + a^2 + \dots + a^{n-1}$.

Proof. Take a prime p dividing n . By hypothesis p divides $a - 1$. It is thus enough to prove that $v_p\left(\frac{a^n - 1}{a - 1}\right) \geq v_p(n)$, which follows from theorem 6.22. \square

The next result, more technical, refines the previous one. One has to be careful when applying this result, since there are a few hypotheses involved in its statement.

Theorem 6.24. (*Lifting the exponent lemma*) Let p be an **odd** prime and let a, b integers not divisible by p such that $p \mid a - b$. Then for all $n \geq 1$

$$v_p(a^n - b^n) = v_p(n) + v_p(a - b).$$

Proof. Call an integer $n \geq 1$ good if satisfies the conclusion of the theorem for any a, b as in the statement. Note that if m, n are good, then so is mn . Indeed, if a, b satisfy the hypotheses of the theorem, then so do a^m and b^m , thus

$$\begin{aligned} v_p(a^{mn} - b^{mn}) &= v_p((a^m)^n - (b^m)^n) = v_p(a^m - b^m) + v_p(n) \\ &= v_p(a - b) + v_p(m) + v_p(n) = v_p(a - b) + v_p(mn) \end{aligned}$$

and mn is good. Since 1 is clearly good, it suffices to prove that any prime q is good. If $q \neq p$, this reduces to proving that $\frac{a^q - b^q}{a - b} = a^{q-1} + a^{q-2}b + \dots + b^{q-1}$ is not divisible by p , which is clear since $a^{q-1} + a^{q-2}b + \dots + b^{q-1} \equiv qa^{q-1} \pmod{p}$ (as $p \mid a - b$) and qa is not divisible by p .

Suppose that $q = p$ and write $a = b + p^k c$ for some integer c not divisible by p and some $k \geq 1$. The binomial formula gives

$$a^p - b^p = p^{k+1}b^{p-1}c + \binom{p}{2}b^{p-2}p^{2k}c^2 + \dots + p^{kp}c^p.$$

Since $p > 2$, the terms $\binom{p}{2}b^{p-2}p^{2k}c, \dots, p^{kp}c^p$ have p -adic valuation greater than $k + 1$, which combined with $\gcd(p, bc) = 1$ gives

$$v_p(a^p - b^p) = v_p(p^{k+1}b^{p-1}c) = k + 1 = 1 + v_p(a - b),$$

as needed. \square

We also mention the following immediate consequence of the previous theorem:

Corollary 6.25. *Let p be an odd prime and let a, b be integers not divisible by p and for which $p \mid a + b$. Then for all **odd** positive integers n*

$$v_p(a^n + b^n) = v_p(a + b) + v_p(n).$$

Proof. It suffices to apply the previous theorem to a and $-b$. \square

The reader might wonder what happens when $p = 2$. In this case the formula is a bit more complicated to state, but much easier to prove.

Theorem 6.26. *If x, y are odd integers and n is an **even** positive integer, then*

$$v_2(x^n - y^n) = v_2\left(\frac{x^2 - y^2}{2}\right) + v_2(n).$$

Proof. Write $n = 2^k a$ for some odd number a . Then using repeatedly the difference of squares formula we obtain

$$x^n - y^n = (x^a - y^a)(x^a + y^a)(x^{2a} + y^{2a}) \dots (x^{2^{k-1}a} + y^{2^{k-1}a}).$$

Observe that if u, v are odd numbers, then $u^2 + v^2 \equiv 2 \pmod{4}$. The previous formula gives therefore

$$v_2(x^n - y^n) = v_2(x^{2a} - y^{2a}) + k - 1.$$

Finally, since a, x, y are odd, it is easy to see that $\frac{x^{2a} - y^{2a}}{x^2 - y^2} = x^{2(a-1)} + \dots + y^{2(a-1)}$ is odd. The result follows. \square

Remark 6.27. When n is odd, things are very simple: $\frac{x^n - y^n}{x - y} = x^{n-1} + \dots + y^{n-1}$ is odd and so $v_2(x^n - y^n) = v_2(x - y)$.

The next series of examples illustrate the power of the previous theorems.

Example 6.28. Find all integers $a, n > 1$ such that any prime factor of $a^n - 1$ is a prime factor of $a - 1$.

Proof. Let p be a prime factor of n and assume that $p > 2$. Any prime factor q of $1 + a + \dots + a^{p-1}$ divides $a^p - 1 \mid a^n - 1$, thus it divides $a - 1$. But then $1 + a + \dots + a^{p-1} \equiv p \pmod{q}$ and since $q \mid 1 + a + \dots + a^{p-1}$ we obtain $q = p$. In other words $1 + a + \dots + a^{p-1} = p^k$ for some $k > 0$, and moreover $p \mid a - 1$. Now lifting the exponent lemma yields (since $p > 2$)

$$v_p(1 + a + \dots + a^{p-1}) = v_p(a^p - 1) - v_p(a - 1) = 1$$

and so $k = 1$. But this is impossible, since $a > 1$ and so $1 + a + \dots + a^{p-1} > p$.

Hence any prime factor p of n is 2, in other words $n = 2^k$ for some $k > 0$. But then $a + 1 \mid a^n - 1$ and so any prime factor of $a + 1$ divides $a - 1$ and so divides 2. Thus $a + 1$ is also a power of 2, say $a + 1 = 2^m$.

Suppose that $k > 1$, i.e. $n > 2$. Then $a^2 + 1 \mid a^n - 1$ and as above we obtain that $a^2 + 1$ is a power of 2, say $a^2 + 1 = 2^l$. Since 4 cannot divide $a^2 + 1$, we must have $l = 2$ and then $a = 1$, a contradiction. Hence $n = 2$ and $a + 1$ is a power of 2. Conversely, if these conditions are satisfied then clearly (a, n) is a solution of the problem. \square

Example 6.29. Find all integers $a, n > 1$ such that any prime factor of $a^n + 1$ is a prime factor of $a + 1$.

Proof. Assume first that n is even. If $p \mid a^n + 1$ is a prime, then $p \mid a + 1$ and so $0 \equiv a^n + 1 \equiv 2 \pmod{p}$, that is $p = 2$. It follows that $a^n + 1 = 2^k$ for some $k > 0$. Since $a^n + 1$ is of the form $x^2 + 1$ (as n is even), and such a number is never a multiple of 4, it follows that $k = 1$, contradicting $a > 1$. Hence n is odd.

If p is a prime factor of n , then any prime q dividing

$$\frac{a^p + 1}{a + 1} = a^{p-1} - a^{p-2} + \dots - a + 1$$

divides $a + 1$ and so $0 \equiv a^{p-1} - a^{p-2} + \dots - a + 1 \equiv p \pmod{q}$, which in turn yields $p = q$. Thus $\frac{a^p+1}{a+1} = p^k$ for some $k > 0$ and using again the lifting the exponent lemma we obtain $k = 1$, thus

$$a^p + 1 = p(a + 1) \quad \text{or} \quad a(a^{p-1} - p) = p - 1.$$

Moreover, from the above discussion we know that $p \mid a + 1$, so $a \geq p - 1$ and the previous equation yields $a^{p-1} - p \leq 1$. Since $p > 2$, we obtain $p + 1 \geq a^{p-1} \geq a^2 \geq (p - 1)^2$, which immediately implies $p = 3$ and then $a = 2$. Hence $a = 2$ and n is a power of 3. If $n \neq 3$, then replacing a with $b = a^{\frac{n}{3}}$ we obtain that any prime factor of $b^3 + 1$ divides $b + 1$ and by the above discussion this forces $b = 2$, which is not the case. Hence $n = 3$ and $a = 2$ is the unique solution of the problem. \square

Remark 6.30. The previous exercise is a generalization of an IMO Shortlist 2000 problem: find all triplets of positive integers (a, m, n) such that $a^m + 1 \mid (a + 1)^n$.

Example 6.31. (IMO Shortlist 1997) Let b, m, n be positive integers such that $b > 1$ and $m \neq n$. Prove that if $b^m - 1$ and $b^n - 1$ have the same prime divisors, then $b + 1$ is a power of 2.

Proof. Without loss of generality we may assume that $m > n$. Let $d = \gcd(m, n)$ and let $m = kd$ and $a = b^d$. Note that $k > 1$ and any prime p dividing $a^k - 1 = b^m - 1$ divides $b^n - 1$ and so it divides $\gcd(b^m - 1, b^n - 1) = b^d - 1 = a - 1$. By example 6.28 we deduce that $a + 1$ is a power of 2, that is $b^d + 1$ is a power of 2. If d is even, then $b^d + 1$ is not a multiple of 4 and is greater than 2, so it cannot be a power of 2. Hence d is odd and this implies that $b + 1$ is a power of 2, since $b + 1 \mid b^d + 1$. \square

Example 6.32. (generalization of IMO 1990 and 1999) Find all primes p and all positive integers n such that n^{p-1} divides $(p - 1)^n + 1$.

Proof. Note that if $p = 2$, then $n = 1$ or $n = 2$. From now on, we assume that $p > 2$. If n is even, then 4 cannot divide n^{p-1} (because 4 does not divide $(p - 1)^n + 1$) and so $p = 2$, a contradiction. So, n is odd. Let q be the smallest

prime factor of n . Since q divides $(p-1)^{2n} - 1$ and $(p-1)^{q-1} - 1$ and since $\gcd(2n, q-1) = 2$, it follows that q divides $(p-1)^2 - 1 = p(p-2)$.

Suppose first that q divides $p-2$. Then, by the lifting the exponent lemma we have

$$(p-1)v_q(n) = v_q(n^{p-1}) \leq v_q((p-1)^{2n} - 1) = v_q((p-1)^2 - 1) + v_q(n),$$

so that $(p-2)v_q(n) \leq v_q(p-2)$. In particular, $p-2 \geq q^{p-2} \geq 3^{p-2}$. This easily implies that $p = 3$, contradicting the fact that q divides $p-2$.

Next, assume that $q = p$, so that again by the lifting exponent lemma (using that n is odd) we have

$$(p-1)v_p(n) = v_p(n^{p-1}) \leq v_p((p-1)^n + 1) = 1 + v_p(n).$$

Thus $(p-2)v_p(n) \leq 1$. In particular, $p = 3$ and $v_p(n) = 1$. Write $n = 3a$ with $\gcd(a, 3) = 1$ and observe that a^2 divides $8^a + 1$. We claim that $a = 1$. Otherwise, let r be the smallest prime factor of a , so that r divides $64^a - 1$ and $64^{r-1} - 1$. Thus r divides 63, since $\gcd(a, r-1) = 1$. But then $r = 3$ or $r = 7$. Since 3 does not divide a , we must have $r = 7$ and 7 divides $8^a + 1$. Since this is of course impossible, it follows that $a = 1$ and $n = 3$. \square

Example 6.33. (China TST 2009) Let n be a positive integer and let $a > b > 1$ be integers such that b is odd and $b^n | a^n - 1$. Prove that $a^b > \frac{3^n}{n}$.

Proof. Take any prime factor p of b , then necessarily $p > 2$ and the lifting the exponent lemma (combined with Fermat's little theorem) gives

$$n \leq v_p(b^n) \leq v_p(a^n - 1) \leq v_p((a^{p-1})^n - 1^n) = v_p(a^{p-1} - 1) + v_p(n),$$

so that

$$a^b > a^{p-1} - 1 \geq p^{v_p(a^{p-1}-1)} \geq \frac{p^n}{n} \geq \frac{3^n}{n}.$$

The result follows. \square

We end this section with the following difficult problem.

Example 6.34. (China TST 2002) Find all positive integers n for which $(2^n - 1)(3^n - 1)$ is a perfect square.

Proof. We will prove that there is no such n . Assume that $(2^n - 1)(3^n - 1) = m^2$ for some integers $m, n \geq 1$. Note that m is even, thus $4 \mid 3^n - 1$ and n is even. Therefore $3 \mid m$ and so $9 \mid 2^n - 1$, which forces $6 \mid n$. Next, we will prove that $10 \mid n$. Write $n = 6k$, thus $(64^k - 1)(3^{6k} - 1) = m^2$ and so

$$(2^k - 1)(16^k - 1) \equiv m^2 \pmod{31}.$$

One easily checks that the left-hand side is a multiple of 31 if and only if $5 \mid k$. Suppose that 5 does not divide k . The previous congruence gives

$$\left(\frac{2^k - 1}{31}\right) \cdot \left(\frac{16^k - 1}{31}\right) = 1,$$

which is equivalent to

$$\left(\frac{2^k + 1}{31}\right) \cdot \left(\frac{4^k + 1}{31}\right) = 1.$$

To check that the last equality is impossible, it suffices to do so for $k = 1, 2, 3, 4$ (using the 31-periodicity of Legendre's symbol modulo 31), which (after simple algebra) comes down to checking the impossibility of any of the following relations

$$\left(\frac{3}{31}\right) \cdot \left(\frac{5}{31}\right) = 1, \quad \left(\frac{5}{31}\right) \cdot \left(\frac{17}{31}\right) = 1, \quad \left(\frac{5}{31}\right) \cdot \left(\frac{13}{31}\right) = 1, \quad \left(\frac{17}{31}\right) = 1.$$

These follows directly from

$$\left(\frac{3}{31}\right) = \left(\frac{17}{31}\right) = \left(\frac{13}{31}\right) = -1, \quad \left(\frac{5}{31}\right) = 1,$$

all easily established.

Write now $n = 10x$ and use the lifting the exponent lemma to obtain

$$\begin{aligned} 2v_{11}(m) &= v_{11}((2^n - 1)(3^n - 1)) = v_{11}((2^{10})^x - 1) + v_{11}((3^{10})^x - 1) \\ &= v_{11}(2^{10} - 1) + v_{11}(3^{10} - 1) + 2v_2(x) = 2v_2(x) + 3, \end{aligned}$$

a contradiction. Thus there are no such n . □

6.2 Legendre's formula

In this section we discuss Legendre's formula giving the p -adic valuation of $n!$ and its consequences to the arithmetic of binomial coefficients. We will use these properties in the next section to obtain nontrivial estimates on the distribution of prime numbers.

6.2.1 The p -adic valuation of $n!$: the exact formula

We have already given several proofs of the fact that the product of n consecutive integers is a multiple of $n!$. Most of these proofs used specific properties of binomial coefficients. We would like to give a proof of this result using the local-global principle according to which $a \mid b$ if and only if $v_p(a) \leq v_p(b)$ for all primes p . For that, it is necessary to compute $v_p(n!)$ for a prime p and a positive integer n . This is the object of the next theorem.

Theorem 6.35. (*Legendre*) *For all primes p and all positive integers n we have*

$$v_p(n!) = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \dots$$

Before giving the proof of this theorem, we emphasize that the apparently infinite sum appearing in the statement is in fact finite, since all but finitely many terms are zero. Indeed, there is k such that $p^k > n$, and then $\left\lfloor \frac{n}{p^i} \right\rfloor = 0$ for all $i \geq k$.

Proof. We have

$$v_p(n!) = v_p(1 \cdot 2 \cdot \dots \cdot n) = v_p(1) + v_p(2) + \dots + v_p(n).$$

Among the numbers $1, 2, \dots, n$ there are $\left\lfloor \frac{n}{p} \right\rfloor$ multiples of p , $\left\lfloor \frac{n}{p^2} \right\rfloor$ multiples of p^2 , and so on. Multiples of p but not of p^2 have contribution 1 to the sum, multiples of p^2 but not of p^3 have contribution 2, and so on. Hence

$$v_p(n!) = \left\lfloor \frac{n}{p} \right\rfloor - \left\lfloor \frac{n}{p^2} \right\rfloor + 2 \left(\left\lfloor \frac{n}{p^2} \right\rfloor - \left\lfloor \frac{n}{p^3} \right\rfloor \right) + 3 \left(\left\lfloor \frac{n}{p^3} \right\rfloor - \left\lfloor \frac{n}{p^4} \right\rfloor \right) + \dots$$

and the sum telescopes to the desired formula. \square

Let us come back now to our original problem, namely giving a proof of the fact that $n!$ divides $(x+1)(x+2)\dots(x+n)$ for any integer x using p -adic valuations. Fix a prime p and let n_k be the number of multiples of p^k among $x+1, \dots, x+n$. As in the proof of the above theorem, we see that

$$v_p((x+1)(x+2)\dots(x+n)) = n_1 + n_2 + \dots$$

On the other hand, it is clear that

$$n_k = \left\lfloor \frac{x+n}{p^k} \right\rfloor - \left\lfloor \frac{x}{p^k} \right\rfloor \geq \left\lfloor \frac{n}{p^k} \right\rfloor,$$

since in general $\lfloor x+y \rfloor \geq \lfloor x \rfloor + \lfloor y \rfloor$ for all real numbers x, y . Thus Legendre's formula yields

$$v_p((x+1)\dots(x+n)) \geq v_p(n!)$$

for all primes p and the result follows.

Here are a few more examples of counting arguments used to establish divisibilities or identities.

Example 6.36. (China TST 2004) Let m_1, m_2, \dots, m_r and n_1, n_2, \dots, n_s be positive integers such that for any integer $d > 1$ the number of multiples of d among m_1, \dots, m_r is greater than or equal to the number of multiples of d among n_1, \dots, n_s . Prove that $n_1 n_2 \dots n_s$ divides $m_1 m_2 \dots m_r$.

Proof. For $d > 1$, let M_d and N_d be the number of multiples of d among m_1, \dots, m_r , respectively n_1, \dots, n_s . By hypothesis $M_d \geq N_d$ for all $d > 1$. For any prime p we have (arguing as in the proof of Legendre's formula)

$$v_p(m_1 m_2 \dots m_r) = M_p + M_{p^2} + \dots + M_{p^n} + \dots \geq N_p + N_{p^2} + \dots = v_p(n_1 n_2 \dots n_s)$$

hence $n_1 \dots n_s \mid m_1 \dots m_r$ and the problem is solved. \square

Example 6.37. (Putnam 2003) Prove that for each positive integer n ,

$$n! = \prod_{i=1}^n \text{lcm}(1, 2, \dots, \lfloor n/i \rfloor).$$

Proof. It is enough to prove that both sides have the same p -adic valuation for all primes p . Fix a prime p . Using Legendre's formula and the fact that

$$v_p(\text{lcm}(1, 2, \dots, d)) = \left\lfloor \log_p(d) \right\rfloor,$$

we are reduced to proving the equality

$$\sum_{k \geq 1} \left\lfloor \frac{n}{p^k} \right\rfloor = \sum_{i=1}^n \left\lfloor \log_p \left\lfloor \frac{n}{i} \right\rfloor \right\rfloor$$

for all primes p and all n . For this, we count in two different ways pairs of positive integers (i, k) such that $ip^k \leq n$. For fixed i there are $\left\lfloor \log_p \left\lfloor \frac{n}{i} \right\rfloor \right\rfloor$ possibilities for k , while for fixed k there are $\left\lfloor \frac{n}{p^k} \right\rfloor$ possibilities for i . The result follows. \square

Example 6.38. (Miklos Schweitzer Competition 1973) Let n, k be positive integers such that $n > k + \text{lcm}(1, 2, \dots, k)$. Prove that $\binom{n}{k}$ has at least k distinct prime factors.

Proof. Write $L_k = \text{lcm}(1, 2, \dots, k)$. It suffices to prove that for $n > k + L_k$ the number $\binom{n}{k}$ is a multiple of a product of k numbers that are pairwise relatively prime and greater than 1. For $0 \leq i < k$ let

$$x_i = \frac{n - i}{\gcd(n - i, L_k)}.$$

Clearly $x_i > 1$ and one easily checks that x_0, \dots, x_{k-1} are pairwise relatively prime (see the proof of example 6.19). It suffices therefore to prove that

$$x_0 x_1 \dots x_{k-1} \mid \binom{n}{k},$$

which is equivalent to

$$k! \mid \prod_{i=n-k+1}^n \gcd(i, L_k).$$

It suffices therefore to prove that for all primes p

$$v_p(k!) \leq \sum_{i=n-k+1}^n v_p(\gcd(i, L_k)).$$

Let $r = v_p(L_k) = \lfloor \log_p k \rfloor$ (see example 6.7). For all $i \leq r$ there are at least $\left\lfloor \frac{k}{p^i} \right\rfloor$ multiples of p^i among $n, n-1, \dots, n-k+1$. Also, if u is a multiple of p^i with $i \leq r$, then so is $\gcd(L_k, u)$. The desired inequality is then an immediate consequence of Legendre's formula. \square

6.2.2 The p -adic valuation of $n!$: inequalities

Observe that for all primes p and all positive integers n we have

$$\left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \dots < \frac{n}{p} + \frac{n}{p^2} + \dots = \frac{n}{p-1} \quad \text{and} \quad \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \dots > \frac{n}{p} - 1.$$

Combining these inequalities with Legendre's formula we obtain the following estimate, which is more useful in many situations than the exact formula for $v_p(n!)$ obtained in the previous section.

Theorem 6.39. *For all $n > 1$ and all primes p we have*

$$\frac{n}{p} - 1 < v_p(n!) < \frac{n}{p-1}.$$

We give now some nice illustrations of the previous result.

Example 6.40. (MEMO 2015) Find all pairs (a, b) of positive integers such that

$$a! + b! = a^b + b^a.$$

Proof. By symmetry, we may assume that $a \leq b$. If $a = 1$, the equation becomes $b! = b$, yielding the solutions $(1, 1)$ and $(1, 2)$, so assume that $a \geq 2$. Then $b! - a^b = b^a - a! \geq a^a - a! > 0$, thus $b! > a^b$. On the other hand, the AM-GM inequality yields

$$b! = 1 \cdot 2 \cdot \dots \cdot b \leq \left(\frac{b(b+1)}{2b} \right)^b = \left(\frac{b+1}{2} \right)^b.$$

We conclude that $2a < b + 1$, thus $b \geq 2a$.

Let p be a prime divisor of a . Then $p \mid a! + b!$ and $p \mid a^b$, thus $p \mid b$. Therefore $v_p(a^b + b^a) \geq a$. On the other hand, since $b \geq 2a$ we have $p \mid (a+1) \cdot (a+2) \cdot \dots \cdot b$, hence

$$v_p(a! + b!) = v_p(a!) + v_p(1 + (a+1) \cdot (a+2) \cdot \dots \cdot b) = v_p(a!) < a,$$

the last inequality being a direct consequence of theorem 6.39. We obtain therefore the plain contradiction $a < a$, showing that all solutions of the problem are $(1, 1)$, $(1, 2)$ and $(2, 1)$. \square

Example 6.41. (Saint Petersburg 2007) Find all positive integers n and k for which

$$1^n + 2^n + \dots + n^n = k!.$$

Proof. We will prove that $n = k = 1$ is the unique solution of the problem. Suppose that $n > 1$. Note that $k^k > k! > n^n$, thus $k > n$. First, assume that n is odd. Then $2^n + 3^n + \dots + n^n$ is a multiple of $n + 2$ (since each of the numbers $2^n + n^n, 3^n + (n-1)^n, \dots$ is a multiple of $n + 2$), thus $k! - 1$ is a multiple of $n + 2$. In particular $k < n + 2$ and since $k > n$ we must have $k = n + 1$. Then $(n + 1)! > n^n$, which gives $n < 3$, a contradiction.

Hence n is even, say $n = 2m$. Also, $4 \mid k!$ and

$$1^n + 2^n + \dots + n^n \equiv m \pmod{4}$$

thus $4 \mid m$ and $8 \mid n$. Write $n = 2^s m$ with $s \geq 3$ and m odd. For $i \in \{1, 2, \dots, n\}$ odd we have

$$i^n = (i^{2^s})^m \equiv 1 \pmod{2^{s+1}}$$

and when i is even $i^n \equiv 0 \pmod{2^{s+1}}$. Thus

$$1^n + 2^n + \dots + n^n \equiv 2^{s-1}m \pmod{2^{s+1}}$$

and so

$$v_2(k!) = v_2(1^n + 2^n + \dots + n^n) = s - 1.$$

On the other hand theorem 6.39 gives

$$v_2(k!) > \frac{k}{2} - 1 > \frac{n}{2} - 1 = 2^{s-1}m - 1 \geq 2^{s-1} - 1,$$

hence $s > 2^{s-1}$, impossible. Hence there are no solutions with $n > 1$. \square

The next example is much more challenging.

Example 6.42. (Russia 2012) Prove that there is a positive integer n such that $1! + 2! + \dots + n!$ has a prime factor greater than 10^{2012} .

Proof. Let $f(n) = 1! + 2! + \dots + n!$ and let S be the set of all primes not exceeding $d := 10^{2012}$. Suppose that for all $n \geq 1$, all prime factors of $f(n)$ are in S . Let $P = \prod_{p \in S} p^2$. The key ingredient is the following result.

Lemma 6.43. *There is a constant $c > 0$ such that for all $p \leq d$ and all $n \geq c$ relatively prime to P*

$$v_p(f(nP - 2)) \leq v_p((nP)!) - 2.$$

Proof. We will prove that for any $p \leq d$, the inequality $v_p(f(nP - 2)) \geq v_p((nP)!) - 1$ can hold for at most one n that is relatively prime to P . Fix $p \leq d$ and suppose that this inequality holds for two integers $n < m$ relatively prime to P . Since

$$v_p((nP - 1)!) = v_p((nP)!) - v_p(nP) = v_p((nP)!) - 2,$$

the strong triangle inequality gives

$$\begin{aligned} v_p(f(mP - 2)) &= v_p((nP - 1)! + f(nP - 2) + (nP)! + \dots + (mP - 2)!) \\ &= v_p((nP - 1)!) = v_p((nP)!) - 2. \end{aligned}$$

On the other hand by assumption

$$v_p(f(mP - 2)) \geq v_p((mP)!) - 1.$$

We deduce that $v_p((nP)!) \geq v_p((mP)!) + 1$, which is obviously impossible. The result follows. \square

Let c be as in the previous lemma. We conclude that for all $n \geq c$ relatively prime to P we have

$$v_p(f(nP - 2)) \leq \frac{nP}{p-1} - 2 < nP.$$

Since all prime factors of $f(nP - 2)$ are less than or equal to d , this forces

$$(nP - 2)! < f(nP - 2) \leq \prod_{p \leq d} p^{nP} < d!^{nP}$$

for all $n \geq c$ relatively prime to P . This is clearly impossible. \square

We also point out the following important consequence of Legendre's formula, which will be very useful in obtaining explicit estimates concerning prime numbers.

Theorem 6.44. *Let $n \geq k \geq 0$ be integers and let p be a prime. Then*

$$p^{v_p\left(\binom{n}{k}\right)} \leq n.$$

In other words, all prime powers dividing $\binom{n}{k}$ are smaller than $n + 1$.

Proof. Legendre's formula gives

$$v_p\left(\binom{n}{k}\right) = v_p(n!) - v_p(k!) - v_p((n-k)!) = \sum_{j \geq 1} \left(\left\lfloor \frac{n}{p^j} \right\rfloor - \left\lfloor \frac{k}{p^j} \right\rfloor - \left\lfloor \frac{n-k}{p^j} \right\rfloor \right).$$

Note that each term in the sum is equal to 0 or 1, since for any real numbers x, y

$$\lfloor x + y \rfloor - \lfloor x \rfloor - \lfloor y \rfloor \in \{0, 1\}.$$

Indeed, the left-hand side equals $\lfloor r + s \rfloor$, where $r = x - \lfloor x \rfloor \in [0, 1)$ and $s = y - \lfloor y \rfloor \in [0, 1)$. Finally note that for $p^j > n$

$$\left\lfloor \frac{n}{p^j} \right\rfloor = \left\lfloor \frac{k}{p^j} \right\rfloor = \left\lfloor \frac{n-k}{p^j} \right\rfloor = 0,$$

thus there are at most $\lfloor \log_p(n) \rfloor$ nonzero terms in the sum and so

$$v_p\left(\binom{n}{k}\right) \leq \lfloor \log_p(n) \rfloor.$$

The result follows. \square

Remark 6.45. The inequality (discussed in the proof of the previous theorem)

$$0 \leq \lfloor x + y \rfloor - \lfloor x \rfloor - \lfloor y \rfloor \leq 1$$

will be used implicitly quite often from now on.

The following example uses similar ideas to establish a rather remarkable identity.

Example 6.46. (AMM E 2686) Let n be an integer greater than 1. Prove that

$$(n+1) \operatorname{lcm} \left(\binom{n}{0}, \binom{n}{1}, \dots, \binom{n}{n} \right) = \operatorname{lcm}(1, 2, \dots, n+1).$$

Proof. We will prove that for each prime p both sides have the same p -adic valuation, which is enough to conclude. Let p be a prime and let k be such that $p^k \leq n+1 < p^{k+1}$. By example 6.7 we have

$$v_p(\operatorname{lcm}(1, 2, \dots, n+1)) = k.$$

Note that $(n+1)\binom{n}{p^{k-1}} = p^k \binom{n+1}{p^k}$, thus the p -adic valuation of the left-hand side is greater than or equal to k . To prove that this valuation is at most k , fix $0 \leq i \leq n$ and use Legendre's formula to get

$$v_p \left((n+1) \binom{n}{i} \right) = v_p \left((i+1) \binom{n+1}{i+1} \right) = v_p(i+1) + \sum_{r \geq 1} x_r,$$

where

$$x_r = \left\lfloor \frac{n+1}{p^r} \right\rfloor - \left\lfloor \frac{i+1}{p^r} \right\rfloor - \left\lfloor \frac{n-i}{p^r} \right\rfloor.$$

Note that $x_r \in \{0, 1\}$ for all r (see remark 6.45) and $x_r = 0$ if $r > k$ (since in this case $p^r > n+1$). The key point is that $x_r = 0$ for all $r \leq v_p(i+1)$. Indeed, writing $i+1 = p^r u$ for some integer u , we have

$$x_r = \left\lfloor \frac{n+1}{p^r} \right\rfloor - u - \left\lfloor \frac{n+1}{p^r} - u \right\rfloor = 0.$$

Putting these observations together yields

$$\sum_{r \geq 1} x_r \leq k - v_p(i+1),$$

from which we get $v_p\left((i+1)\binom{n+1}{i+1}\right) \leq k$ for all $0 \leq i \leq n$, establishing the desired inequality. \square

Combining the result of the previous exercise with example 6.8 yields the following estimate for the number $\pi(n)$ of primes not exceeding n , which is surprisingly good (see the next section for a more detailed discussion of such issues).

Example 6.47. Prove that for all $n > 1$ we have

$$\text{lcm}(1, 2, \dots, n) \geq 2^{n-1} \quad \text{and} \quad n^{\pi(n)} \geq 2^{n-1}.$$

Proof. For the first inequality, simply note that

$$(n+1) \text{lcm}\left(\binom{n}{0}, \binom{n}{1}, \dots, \binom{n}{n}\right) \geq \sum_{j=0}^n \binom{n}{j} = 2^n$$

and use the result established in the previous example. For the second inequality, use the first one and the inequality $\text{lcm}(1, 2, \dots, n) \leq n^{\pi(n)}$ established in example 6.8. \square

Example 6.48. Prove that if $c \in (0, 2)$, then for all sufficiently large n the product of all primes not exceeding n is greater than c^n .

Proof. By the previous example

$$\text{lcm}(1, 2, \dots, n) \geq 2^{n-1}.$$

On the other hand by example 6.7 we have

$$\text{lcm}(1, 2, \dots, n) = \prod_{p \leq n} p^{\lfloor \log_p(n) \rfloor} \leq \prod_{p \leq \sqrt{n}} n \cdot \prod_{\sqrt{n} < p \leq n} p \leq n^{\sqrt{n}} \cdot \prod_{p \leq n} p.$$

We deduce that

$$\prod_{p \leq n} p \geq 2^{n-1} \cdot n^{-\sqrt{n}}.$$

Thus we need to prove that for any $c \in (0, 2)$ we have

$$\left(\frac{2}{c}\right)^n \geq 2n^{\sqrt{n}}$$

for all n large enough. Since $2n^{\sqrt{n}} < n^{2\sqrt{n}}$, it suffices to check that $\left(\frac{2}{c}\right)^{\sqrt{n}} > n^2$ for large enough n , which is immediate. \square

6.2.3 Kummer's theorem

Instead of giving estimates for $\left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \dots$ like we did in theorem 6.39, we can also obtain an exact formula as follows: write

$$n = a_k p^k + a_{k-1} p^{k-1} + \dots + a_0$$

in base p (thus $a_0, \dots, a_k \in \{0, 1, \dots, p-1\}$ and $a_k \neq 0$). Then for all $0 \leq j \leq k$

$$\left\lfloor \frac{n}{p^j} \right\rfloor = a_k p^{k-j} + a_{k-1} p^{k-1-j} + \dots + a_j,$$

therefore

$$\begin{aligned} \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \dots &= \sum_{j=1}^k (a_k p^{k-j} + a_{k-1} p^{k-1-j} + \dots + a_j) \\ &= a_k (p^{k-1} + p^{k-2} + \dots + 1) + a_{k-1} (p^{k-2} + p^{k-3} + \dots + 1) + \dots + a_1 \\ &= a_k \cdot \frac{p^k - 1}{p - 1} + a_{k-1} \frac{p^{k-1} - 1}{p - 1} + \dots + a_1 \frac{p - 1}{p - 1} \\ &= \frac{(a_k p^k + \dots + a_1 p + a_0) - (a_k + \dots + a_0)}{p - 1} = \frac{n - s_p(n)}{p - 1}, \end{aligned}$$

where

$$s_p(n) = a_0 + \dots + a_k$$

is the sum of digits of n when written in base p . Combining this computation with Legendre's theorem we obtain the following result.

Theorem 6.49. *For all $n \geq 1$ and all primes p we have*

$$v_p(n!) = \frac{n - s_p(n)}{p - 1},$$

where $s_p(n)$ is the sum of digits of n when written in base p .

This theorem immediately implies the following formula for the p -adic valuation of binomial coefficients.

Corollary 6.50. *For all primes p and all integers $n \geq k \geq 1$*

$$v_p \left(\binom{n}{k} \right) = \frac{s_p(k) + s_p(n-k) - s_p(n)}{p-1}.$$

Let us observe that $\frac{s_p(k) + s_p(n-k) - s_p(n)}{p-1}$ is precisely the number of carries when adding k and $n-k$ in base p . We obtain therefore the following beautiful theorem.

Theorem 6.51. (*Kummer*) *The p -adic valuation of $\binom{n}{k}$ is the number of carries when adding k and $n-k$ in base p .*

Remark 6.52. Even more precisely, for each $j \geq 1$ we have

$$\left\lfloor \frac{n}{p^j} \right\rfloor - \left\lfloor \frac{k}{p^j} \right\rfloor - \left\lfloor \frac{n-k}{p^j} \right\rfloor = \frac{u+v-w}{p^j},$$

where u, v, w are the remainders of $k, n-k, n$ when divided by p^j . Note that $u+v=w$ if and only if $u+v < p^j$, if and only if there is no carry in the j th digit when we add k and $n-k$ in base p . Thus

$$\left\lfloor \frac{n}{p^j} \right\rfloor - \left\lfloor \frac{k}{p^j} \right\rfloor - \left\lfloor \frac{n-k}{p^j} \right\rfloor$$

is equal to 1 if there is a carry in the j th digit when adding k and $n-k$ in base p , and it is equal to 0 otherwise.

We illustrate the previous results with some concrete examples.

Example 6.53. Prove that if n is a positive integer and $1 \leq k \leq 2^n$, then

$$v_2 \left(\binom{2^n}{k} \right) = n - v_2(k).$$

Proof. Using corollary 6.50 we obtain

$$v_2 \left(\binom{2^n}{k} \right) = s_2(k) + s_2(2^n - k) - s_2(2^n).$$

If $k = 2^r s$ with $r \geq 0$ and s odd, then clearly $r \leq n$ and

$$s_2(2^n - k) = s_2(2^n - 2^r s) = s_2(2^{n-r} - s) = n - r + 1 - s_2(s) = n - r + 1 - s_2(k).$$

Taking into account that $s_2(2^n) = 1$, the result follows. \square

Example 6.54. Prove that $n \geq 1$ is a power of 2 if and only if 4 does not divide $\binom{2n}{n}$.

Proof. 4 does not divide $\binom{2n}{n}$ if and only if $v_2(\binom{2n}{n}) \leq 1$. This is equivalent to $2s_2(n) - s_2(2n) \leq 1$. Since $s_2(2n) = s_2(n)$ (as the binary expression of $2n$ is simply the binary expression of n followed by a terminal 0), this is further equivalent to $s_2(n) \leq 1$. Clearly, this happens if and only if n is a power of 2. \square

Example 6.55. Prove that all numbers $\binom{2^n}{k}$ with $1 \leq k < 2^n$ are even and exactly one of them is not a multiple of 4. Which one?

Proof. Corollary 6.50 gives

$$v_2 \left(\binom{2^n}{k} \right) = s_2(k) + s_2(2^n - k) - 1 \geq 1.$$

In order to have equality we need $s_2(k) = s_2(2^n - k) = 1$, which is easily seen to happen only for $k = 2^{n-1}$. \square

Example 6.56. (IMO Shortlist 2008) Let n be a positive integer. Prove that the remainders of the numbers

$$\binom{2^n - 1}{0}, \binom{2^n - 1}{1}, \binom{2^n - 1}{2}, \dots, \binom{2^n - 1}{2^{n-1} - 1}$$

when divided by 2^n are a permutation of $1, 3, 5, \dots, 2^n - 1$.

Proof. By Lucas' theorem (or by example 6.53 and the equality $\binom{2^n-1}{k} = \frac{2^n-k}{2^n} \binom{2^n}{k}$) all remainders belong to $\{1, 3, 5, \dots, 2^n - 1\}$, so it suffices to prove that $\binom{2^n-1}{k}$ and $\binom{2^n-1}{l}$ are not congruent modulo 2^n if $1 \leq k < l \leq 2^n$ are odd. Assume that

$$\binom{2^n-1}{k} \equiv \binom{2^n-1}{l} \pmod{2^n}$$

and observe that

$$\begin{aligned} \binom{2^n-1}{l} &= \binom{2^n}{l} - \binom{2^n-1}{l-1} = \binom{2^n}{l} - \binom{2^n}{l-1} + \binom{2^n-1}{l-2} = \dots \\ &= \binom{2^n}{l} - \binom{2^n}{l-1} + \binom{2^n}{l-2} - \dots + \binom{2^n-1}{k}, \end{aligned}$$

thus the congruence can be written as

$$\binom{2^n}{l} - \binom{2^n}{l-1} + \dots - \binom{2^n}{k+1} \equiv 0 \pmod{2^n}.$$

Since $\binom{2^n}{s}$ is divisible by 2^n whenever s is odd (by example 6.53), the previous congruence is equivalent to

$$\binom{2^n}{l-1} + \binom{2^n}{l-3} + \dots + \binom{2^n}{k+1} \equiv 0 \pmod{2^n}.$$

Let

$$N = \min_{s \in \{l-1, l-3, \dots, k+1\}} v_2 \left(\binom{2^n}{s} \right) = v_2 \left(\binom{2^n}{x} \right)$$

for some $x \in \{l-1, l-3, \dots, k+1\}$. Since $\binom{2^n}{l-1} + \binom{2^n}{l-3} + \dots + \binom{2^n}{k+1}$ is a multiple of 2^n and $n > N$, there must be $y \in \{l-1, l-3, \dots, k+1\}$ different from x such that

$$v_2 \left(\binom{2^n}{x} \right) = v_2 \left(\binom{2^n}{y} \right).$$

Using again example 6.53, we obtain $v_2(x) = v_2(y)$. Let $m = v_2(x)$ and without loss of generality, assume that $x < y$. Then $x = 2^m a$ and $y = 2^m b$

with a, b odd and $a < b$. But then $x + 2^m \in \{l - 1, l - 3, \dots, k + 1\}$ and (using once more example 6.53)

$$v_2 \left(\binom{2^n}{x + 2^m} \right) = n - v_2(x + 2^m) \leq n - (m + 1) < n - m = v_2 \left(\binom{2^n}{x} \right),$$

contradicting the minimality property of x . \square

6.3 Estimates for binomial coefficients and the distribution of prime numbers

This section is rather technical, but contains many beautiful results concerning the distribution of prime numbers. The reader may want to skip some of the more involved estimates for a first reading. Our goal is to use Legendre's formula and a detailed study of binomial coefficients and their p -adic valuations to try to answer the following basic question: about how many primes are there between 1 and n ?

6.3.1 Central binomial coefficients and Erdős' inequality

We will focus on central binomial coefficients, since these are the easiest to estimate asymptotically. More precisely, since $\binom{2n}{n}$ is the largest among $\binom{2n}{0}, \dots, \binom{2n}{2n}$ and the sum of these binomial coefficients is 2^{2n} , it is clear that

$$4^n > \binom{2n}{n} \geq \frac{4^n}{2n + 1}.$$

Also note that since $\binom{2n+1}{n} = \binom{2n+1}{n+1}$ and $\sum_{k=0}^{2n+1} \binom{2n+1}{k} = 2^{2n+1}$, we have

$$\binom{2n+1}{n} < 4^n.$$

This will play a crucial role in the proof of the following beautiful result. If S is a set of positive integers, we make the convention that $\prod_{p \in S} p$ is the product of all primes in S (the letter p will always denote a prime in this section).

Theorem 6.57. (Erdős) For $n \geq 2$ the product of all primes not exceeding n is smaller than 4^{n-1} . In other words

$$\prod_{p \leq n} p < 4^{n-1}.$$

Proof. The proof is by strong induction, the case $n = 2$ being clear. Assume that the result holds up to $n - 1$ and let us prove it for $n > 2$. If n is even, then clearly $\prod_{p \leq n} p = \prod_{p \leq n-1} p$ and we are done thanks to the inductive hypothesis. Assume that $n = 2k + 1$ is odd. Note that

$$\binom{2k+1}{k} = \frac{(2k+1)!}{k!(k+1)!} = \frac{(k+2)(k+3)\dots(2k+1)}{k!}$$

is a multiple of $\prod_{k+2 \leq p \leq 2k+1} p$, thus $\prod_{k+2 \leq p \leq 2k+1} p \leq \binom{2k+1}{k}$ and so

$$\prod_{p \leq n} p \leq \prod_{p \leq k+1} p \cdot \binom{2k+1}{k}.$$

By the inductive hypothesis $\prod_{p \leq k+1} p < 4^k$ and by the discussion preceding the theorem $\binom{2k+1}{k} < 4^k$, hence

$$\prod_{p \leq n} p < 4^k \cdot 4^k = 4^{n-1},$$

finishing the proof. □

Example 6.58. Prove that for all sufficiently large integers n there are $2n$ consecutive composite numbers smaller than $n!$.

Proof. Let p_1, \dots, p_k be all primes not exceeding $2n + 1$. Then $p_1 \dots p_k + 2, p_1 \dots p_k + 3, \dots, p_1 \dots p_k + 2n + 1$ are all composite and the largest of these numbers is (by theorem 6.57)

$$p_1 \dots p_k + 2n + 1 < 4^n + 2n + 1 < 2 \cdot 4^n.$$

Since $2 \cdot 4^n < n!$ for n large enough, we are done. □

Example 6.59. Prove that for all $n > 2$ we have

$$\text{lcm}(1, 2, \dots, n) < 9^n.$$

Proof. Combining example 6.7 and theorem 6.57 gives

$$\text{lcm}(1, 2, \dots, n) = \prod_{p \leq n} p^{\lfloor \log_p n \rfloor} = \prod_{p > \sqrt{n}} p \cdot \prod_{p \leq \sqrt{n}} p^{\lfloor \log_p n \rfloor} < 4^n \cdot \left(\prod_{p \leq \sqrt{n}} n \right) \leq 4^n \cdot n^{\sqrt{n}}.$$

It suffices therefore to prove that $4^n \cdot n^{\sqrt{n}} < 9^n$, or equivalently that $\frac{\ln n}{\sqrt{n}} < \ln \frac{9}{4}$. A simple study of the function $f(x) = \frac{\ln x}{\sqrt{x}}$ shows that f is maximal at $x = e^2$ and $f(e^2) = \frac{2}{e} < 0.74 < \ln \frac{9}{4}$. \square

We give now a different and much more conceptual proof of the result established in example 3.31.

Example 6.60. (IMC 2012) Is the set of positive integers n such that $n! + 1$ divides $(2012n)!$ finite or infinite?

Proof. We will prove that there are only finitely many such n . Suppose that $n! + 1$ divides $(kn)!$, where $k = 2012$. Then any prime factor of $n! + 1$ is greater than n and smaller than or equal to kn . If p is such a prime factor, theorem 6.39 combined with the inequality $p > n$ yields

$$v_p(n! + 1) \leq v_p((kn)!) < \frac{kn}{p-1} \leq k.$$

Using theorem 6.57, it follows that

$$n! + 1 = \prod_{n < p \leq kn} p^{v_p(n!+1)} < \prod_{n < p \leq kn} p^k < \left(\prod_{p \leq kn} p \right)^k < 4^{k^2 n}.$$

Thus any solution n of the problem satisfies $n! < 4^{k^2 n}$. It follows immediately that there are only finitely many solutions. \square

6.3.2 Estimating $\pi(n)$

Recall that

$$\pi(n) = \sum_{p \leq n} 1$$

denotes the number of prime numbers not exceeding n . One of the deepest and most beautiful theorems in number theory is the following result proved by Hadamard and de la Vallée-Poussin in 1896. The proof of this result is way beyond the scope of this modest book.

Theorem 6.61. (*prime number theorem*) *We have*

$$\lim_{n \rightarrow \infty} \frac{\pi(n)}{\frac{n}{\ln n}} = 1.$$

The famous prime number theorem asserts that for n large enough $\pi(n)$ behaves like $\frac{n}{\log n}$. The following result gives a **uniform** upper bound for the quotient $\frac{\pi(n)}{\frac{n}{\ln n}}$. Of course, this bound is weaker than the one given by the prime number theorem, but it is rather amazing that with so few tools it already gives the "correct" upper bound. Note that $6 \ln 2 = 4.15\dots$

Theorem 6.62. *For all $n \geq 2$ we have*

$$n^{\pi(n)} < 64^n, \quad \text{or equivalently} \quad \pi(n) < 6 \ln 2 \cdot \frac{n}{\ln n}.$$

Proof. Since

$$\binom{2n}{n} = \frac{(n+1)(n+2)\dots(2n)}{n!}$$

is a multiple of $\prod_{n < p \leq 2n} p$, we deduce that

$$n^{\pi(2n) - \pi(n)} = \prod_{n < p \leq 2n} n < \prod_{n < p \leq 2n} p \leq \binom{2n}{n} \leq 4^n.$$

Setting $n = 2^k$ yields

$$k(\pi(2^{k+1}) - \pi(2^k)) \leq 2^{k+1}, \quad \text{or} \quad (k+1)\pi(2^{k+1}) - k\pi(2^k) \leq 2^{k+1} + \pi(2^{k+1}).$$

Since $\pi(2^{k+1}) \leq 2^k$, we obtain

$$(k+1)\pi(2^{k+1}) - k\pi(2^k) \leq 3 \cdot 2^k.$$

Adding these inequalities for $k = 1, 2, \dots, n-1$ we obtain the inequality

$$n \cdot \pi(2^n) < 3 \cdot 2^n.$$

In general, let $k = \lfloor \log_2(n) \rfloor$, so that $2^k \leq n < 2^{k+1}$. Then using the previously established inequality, we obtain

$$n^{\pi(n)} < (2^{k+1})^{\pi(2^{k+1})} < 8^{2^{k+1}} \leq 64^n,$$

yielding the desired result. \square

We would like to find a good lower bound for $\pi(n)$. Actually we have already obtained a fairly good such bound in the previous section. More precisely we proved the inequality

$$n^{\pi(n)} \geq 2^{n-1}$$

for all $n > 1$ in example 6.47. This can be rewritten as

$$\pi(n) \geq \ln 2 \cdot \frac{n-1}{\ln n},$$

and is a fairly good lower bound taking into account that $\ln 2 = 0.69\dots$ and that $\frac{n-1}{\ln n}$ is essentially the same as $\frac{n}{\ln n}$. In particular, this bound immediately implies the following one, which is weaker but has a somewhat more conceptual proof.

Theorem 6.63. *For $n \geq 2$ we have*

$$n^{\pi(n)} \geq \sqrt{2}^n, \quad \text{or equivalently} \quad \pi(n) \geq \frac{\ln 2}{2} \cdot \frac{n}{\ln n}.$$

Proof. One easily checks the result for $n \leq 5$, so assume that $n > 5$. Writing $n = 2k$ or $n = 2k-1$ and using that $\pi(2k-1) = \pi(2k)$ for $k \geq 2$, it suffices to prove that $(2k-1)^{\pi(2k-1)} \geq 2^k$ for $k \geq 3$. Theorem 6.44 shows that for all

primes p dividing $\binom{2k}{k}$ we have $p^{v_p(\binom{2k}{k})} \leq 2k-1$ (the equality $p^{v_p(\binom{2k}{k})} = 2k$ is impossible, as this would force $p = 2$ and $k = 2^j$ for some j and then $2 = 2k$). Thus

$$\binom{2k}{k} = \prod_{p \leq 2k-1} p^{v_p(\binom{2k}{k})} \leq (2k-1)^{\pi(2k-1)}.$$

Since $\binom{2k}{k} \geq \frac{4^k}{2k+1}$, it suffices to prove that $2^k \geq 2k+1$ for $k \geq 3$, which is immediate. \square

Example 6.64. Prove that for all $n > 1$ we have

$$\frac{n \ln n}{5} < p_n < 6n \ln n.$$

Proof. The key point is that $\pi(p_n) = n$, so we can use the previous estimates. For instance, theorem 6.62 yields

$$64^{p_n} > p_n^n > n^n,$$

thus

$$p_n > \frac{n \ln n}{\ln 64} > \frac{n \ln n}{5}.$$

Similarly, theorem 6.63 yields

$$n \geq \frac{\ln 2}{2} \cdot \frac{p_n}{\ln p_n}.$$

The function $f(x) = \frac{x}{\ln x}$ being increasing for $x \geq 3$ (as a simple derivative computation shows), assuming that $p_n \geq 6n \ln n$ we obtain

$$n \geq \frac{\ln 2}{2} \cdot \frac{6n \ln n}{\ln(6n \ln n)},$$

which yields

$$\ln(6n \ln n) \geq 3 \ln 2 \cdot \ln n > 2 \ln n = \ln n^2.$$

We deduce that $6 \ln n > n$, which is false for $n > 20$ (as one can easily check). For $n \leq 20$ it is not difficult to check the result by hand (taking into account that $p_{20} = 71$). \square

Remark 6.65. Deep theorems of Rosser and Schoenfeld show that if p_n is the n th prime, then $p_n > n \log n$ and for all $n > 66$

$$\frac{n}{\log n - \frac{1}{2}} < \pi(n) < \frac{n}{\log n - \frac{3}{2}}.$$

We illustrate the previous theorems with two beautiful examples.

Example 6.66. Let k be a positive integer. Prove that there is a positive integer n which can be written as the sum of two primes in more than k different ways.

Proof. There are $\pi(N)^2$ pairs of prime numbers (p, q) with $p, q \leq N$. For any such pair the sum $p + q$ is at most $2N$. Therefore by the pigeonhole principle there must be an $r \leq 2N$ which can be written as $r = p + q$ for at least

$$\frac{\pi(N)^2}{2N} \geq \frac{(\ln 2)^2}{4} \cdot \frac{N}{(\ln N)^2}$$

pairs (p, q) (using theorem 6.63). This quantity tends to infinity as N grows, so for N large enough this implies that r can be written as a sum of primes in at least k ways. \square

Example 6.67. Prove that $\pi(n)$ divides n for infinitely many n .

Proof. The solution of this problem is short, but not easy to find! We claim that for any positive integer $m \geq 2$ we can find an integer n such that $m\pi(n) = n$. We will choose $n = mk$ for some positive integer k , so the previous equation becomes $\frac{\pi(mk)}{mk} = \frac{1}{m}$. Consider the set

$$S = \left\{ j \geq 1 \mid \frac{\pi(mj)}{mj} \geq \frac{1}{m} \right\}.$$

Note that $1 \in S$, so S is nonempty. Since $\frac{\pi(x)}{x}$ tends to 0 as $x \rightarrow \infty$, the set S is finite. Letting $k = \max(S)$, we will prove that $\frac{\pi(mk)}{mk} = \frac{1}{m}$, which will finish the proof. If $\frac{\pi(mk)}{mk} > \frac{1}{m}$ does not hold, then $\pi(m(k+1)) \geq \pi(mk) \geq k+1$, contradicting the maximality of k . The result follows. \square

6.3.3 Bertrand's postulate

The last result we want to establish in this section is the following theorem, that was conjectured by Bertrand in 1845 and proved by Chebyshev in 1850. Later on, Erdős simplified the proof, and we follow his approach here. The proof is unfortunately fairly technical and we advise the reader to skip it for a first reading.

Theorem 6.68. (*Bertrand's postulate*) *For all $n \geq 4$ there is a prime $p \in (n, 2n - 2)$. In particular, for $n > 1$ there is always a prime between n and $2n$.*

The key of the proof is again the study of the prime factorization of $\binom{2n}{n}$. It will be useful to introduce the following expression

$$P_n = \prod_{n < p \leq 2n-1} p,$$

the product of all primes between n and $2n$. Since it is not at all clear that there are such primes (this is after all what we are trying to prove!), we use the convention that $P_n = 1$ if there are no such primes. We will actually prove a much stronger result (see the discussion following the proof of the next theorem for the reason why it is much stronger than Bertrand's postulate).

Theorem 6.69. *For all $n > 125$ we have*

$$P_n > \frac{4^{\frac{n}{3}}}{(2n)\sqrt{\frac{n}{2}}}.$$

Proof. Let $A = \binom{2n}{n}$. All prime factors of A are between 1 and $2n$ and it is a simple matter to check the equality

$$A = P_n \cdot \prod_{p \leq n} p^{v_p(A)}.$$

Note that

$$A \geq \frac{4^n}{2n+1} > \frac{4^{n-1}}{2n},$$

thus in order to prove the theorem it suffices to prove that

$$\prod_{p \leq n} p^{v_p(A)} < (2n)^{\sqrt{\frac{n}{2}}-1} \cdot 4^{\frac{2n}{3}-1}.$$

For this, we will carefully analyze each $p^{v_p(A)}$. By theorem 6.44 each $p^{v_p(A)}$ is $\leq 2n$. Also, Legendre's formula shows that $v_p(A) \leq 1$ for $p > \sqrt{2n}$ and, most importantly, that $v_p(A) = 0$ for $p \in (2n/3, n]$. Indeed, for such p we have $v_p((2n)!) = 2$ and $v_p(n!) = 1$, thus $v_p(A) = 0$. We conclude that

$$\prod_{p \leq n} p^{v_p(A)} \leq \prod_{p \leq \sqrt{2n}} (2n) \cdot \prod_{\sqrt{2n} < p \leq \frac{2n}{3}} p.$$

Now let $n \geq 125$ and let $k = \lfloor \sqrt{2n} \rfloor$, so that $k \geq 15$. Since $1, 9, 15, 4, \dots, 2 \lfloor \frac{k}{2} \rfloor$ are not primes, we have

$$\pi(k) \leq k - (2 + \lfloor \frac{k}{2} \rfloor) < \frac{k}{2} - 1 \leq \sqrt{\frac{n}{2}} - 1.$$

Combining these observations with theorem 6.57 finally yields

$$\prod_{p \leq n} p^{v_p(A)} < (2n)^{\sqrt{\frac{n}{2}}-1} \cdot 4^{\frac{2n}{3}-1},$$

as needed. □

This fairly technical statement hides quite a lot of interesting information. For instance, since we trivially have

$$P_n < (2n)^{\pi(2n)-\pi(n)},$$

the previous theorem yields

$$\pi(2n) - \pi(n) > \frac{\ln 4}{3} \cdot \frac{n}{\ln 2n} - \sqrt{\frac{n}{2}},$$

which shows that given $c > \frac{\ln 4}{3}$, we have

$$\pi(2n) - \pi(n) > c \frac{n}{\ln n}$$

for all sufficiently large n , in other words there are many primes between n and $2n$ for n large enough.

We still have to explain why theorem 6.69 implies Bertrand's postulate. We assume from now on that $n > 225$ (using tables of primes, one checks that Bertrand's postulate holds up to 225). Assume that there is no prime $p \in (n, 2n - 2)$. This means that in the product defining P_n there can be at most one term, namely $2n - 1$, in particular $P_n \leq 2n - 1 < 2n$. Using theorem 6.69 we obtain the inequality

$$4^{\frac{n}{3}} < (2n)^{1+\sqrt{\frac{n}{2}}} < (2n)^{\sqrt{n}}$$

and so

$$2^{\frac{\sqrt{n}}{3}} < \sqrt{2} \cdot \sqrt{n}.$$

Letting $k = \left\lfloor \frac{\sqrt{n}}{3} \right\rfloor$, we have $k \geq 5$ and the previous inequality yields

$$2^k < 3 \cdot \sqrt{2} \cdot (k+1) < 5(k+1).$$

It is however easy to check by induction that $2^s \geq 5(s+1)$ for $s \geq 5$, yielding the desired contradiction. Note that this argument also shows that there are at least two primes between n and $2n$ for $n > 225$ (one actually checks that this holds for all $n > 5$).

Remark 6.70. a) It is of course not necessary to check that Bertrand's postulate for each $n \leq 224$ in order to finish the proof. Actually, using the sequence of primes

$$7, 11, 13, 19, 23, 37, 43, 73, 83, 139, 163, 277,$$

the postulate is proved in no time at all for $n \leq 225$.

b) Sylvester and Schur proved the following beautiful generalization of Bertrand's postulate: if $n > k$, then at least one of the numbers $n, n+1, \dots, n+k-1$ has a prime factor greater than k . In other words, for $n \geq 2k$ the binomial coefficient $\binom{n}{k}$ has a prime factor greater than k . Erdős proved that for $k \geq 202$ and $n \geq 2k$ we have

$$\binom{n}{k} > n^{\pi(k)},$$

which immediately implies the previous result for such n and k . The proof is unfortunately more technical than that of Bertrand's postulate, even though the key ideas are the same.

c) By a deep theorem of Polya, if $k \geq 2$ is an integer and if $a_1 < a_2 < \dots$ is the sequence of integers all of whose prime factors do not exceed k , then $a_{i+1} - a_i$ tends to ∞ . In particular, if n is large enough, then every integer among $n, n+1, \dots, n+k-1$, with one possible exception, has a prime divisor greater than k .

d) Legendre conjectured that for all sufficiently large n there is a prime between n and $n + \sqrt{n}$. This is still wide open.

After the previous hard work, it is time to see some concrete illustrations of these results. Unfortunately, there seems to be no easier proof for the following one.

Example 6.71. For $n > 1$, $n!$ is not a perfect power.

Proof. We can assume that $n > 3$. By Bertrand's postulate there is a prime between $n/2$ and n . Clearly $v_p(n!) = 1$ and the result follows. \square

Remark 6.72. A difficult theorem of Erdős and Selfridge states that the product of consecutive integers is never a perfect power. The proof is much harder than that of the previous corollary. They actually prove that for all integers $l, k > 1$ and $m \geq 1$ there is a prime $p > k$ whose exponent in $(m+1)(m+2)\dots(m+k)$ is not a multiple of l . Moreover, they conjecture that if $l \geq 2$ and $k \geq 3$ then we can even find such $p > k$ with exponent 1, except in one case, namely for $48 \cdot 49 \cdot 50$ (for $k = 2$ there are infinitely many exceptions).

Example 6.73. Prove that if $n > 1$ then we can make n pairs $(a_1, b_1), \dots, (a_n, b_n)$ out of the numbers $1, 2, \dots, 2n$, such that $a_i + b_i$ is a prime for all $1 \leq i \leq n$.

Proof. We prove this statement by strong induction on n , the case $n = 2$ being clear (consider the groups $(1, 4)$ and $(2, 3)$). Suppose that the statement is true for $n < k$ and let us prove it for $n = k$. By Bertrand's postulate there is a prime p such that $2k > p - 2k \geq 1$. Considering the pairs $(2k, p - 2k)$, $(2k - 1, p - 2k + 1), \dots, \left(\frac{p-1}{2}, \frac{p+1}{2}\right)$ and applying the inductive hypothesis to

$1, 2, \dots, p - 2k - 1$ (note that $p - 2k - 1$ is even and less than $2k$) yields the desired result. \square

Example 6.74. Let A be a subset of $\{1, 2, 3, \dots, 2n\}$ with more than n elements. Prove that there are two distinct elements of A whose sum is a prime number.

Proof. Consider a partition on $\{1, 2, \dots, 2n\}$ into pairs (a_i, b_i) such that $a_i + b_i$ is a prime for all $1 \leq i \leq n$. Since $|A| > n$, there is i such that $a_i, b_i \in A$ and we are done. \square

Example 6.75. Find all disjoint and nonempty subsets $A, B \subset \mathbb{N}$ such that $A \cup B = \mathbb{N}$ and whenever x, y are distinct positive integers belonging simultaneously to A or to B , $x + y$ is composite.

Proof. Clearly letting A be the set of positive even integers and B the set of positive odd integers yields a solution of the problem. We obtain another solution by permuting the role of even and odd numbers. We will prove that there is no other solution. By symmetry we may assume that $1 \in A$, then clearly $2 \in B$ and so $3 \in A$ and $4 \in B$. Suppose now that $n \geq 2$ and that $1, 3, \dots, 2n - 1 \in A$, while $2, 4, \dots, 2n \in B$. By Bertrand's postulate there is a prime $p \in (2n + 1, 2(2n + 1) - 2)$ and then $p - (2n + 1) \in \{2, 4, \dots, 2n\} \subset B$. Using the hypothesis of the problem, it follows that $2n + 1 \in A$. Similarly, considering a prime $p \in (2n + 2, 4n + 2)$ shows that $2n + 2 \in B$. We have just proved by induction that A contains all odd positive integers and that B contains all even integers. The result follows. \square

Example 6.76. (USAMO 2012) For which integers $n > 1$ is there an infinite sequence a_1, a_2, a_3, \dots of nonzero integers such that for all positive integers k

$$a_k + 2a_{2k} + \dots + na_{nk} = 0?$$

Proof. Observe that $n = 2$ is not a solution of the problem. Indeed, the relation $a_k + 2a_{2k} = 0$ for all k forces $2^j \mid a_k$ for all j and k , thus $a_k = 0$ for all k . We will prove that all numbers different from 2 are solutions, by constructing such a sequence. We will moreover impose that $a_m a_n = a_{mn}$ for all positive integers m, n , in particular $a_1 = 1$. Thus we only need to define a_p for all primes p , and moreover the relation $a_k + 2a_{2k} + \dots + na_{nk} = 0$ is

then equivalent to $a_1 + 2a_2 + \dots + na_n = 0$. For $n = 4$ one can define $a_2 = -1$, $a_3 = -1$ and give arbitrary nonzero values to a_p for any prime $p \neq 2, 3$.

Assuming that $n \neq 2, 4$, we will prove in the next paragraph that we can find different primes p, q such that $\sqrt{n} < p \leq n$ and $\frac{n}{2} < q < n$. For any prime r different from p and q define $a_r = q$. Then a_k is a multiple of q for any $k \in \{1, 2, \dots, n\}$ different from $1, p, q$ since any such q has a prime divisor different from p and q (since $p, q > \sqrt{n}$). We only need to give values to a_p and a_q such that

$$\sum_{i=1}^{q-1} ia_i + qa_q + \sum_{i=q+1}^n ia_i = 0,$$

in other words we need to find a value for a_p such that q divides $\sum_{i=1}^{q-1} ia_i + \sum_{i=q+1}^n ia_i$. As we have already observed, this sum is congruent modulo q to $1 + pa_p$, thus we can take any number m for which $q \mid 1 + pm$ and set $a_p = m$. We still need to prove the existence of p and q as above. We will assume that $n \geq 16$, for the other cases it is fairly easy to find explicitly p and q as desired. Applying Bertrand's postulate we can find a prime $q \in (\lfloor \frac{n}{2} \rfloor, 2 \lfloor \frac{n}{2} \rfloor - 2)$. Then $\frac{n}{2} < q < n$. Applying again Bertrand's postulate, there is a prime $p \in (\frac{q}{2}, q)$. Then $p > \frac{q}{2} > \frac{n}{4} \geq \sqrt{n}$ and the claim is proved. \square

Example 6.77. A polynomial $f \in \mathbb{Z}[X, Y]$ with integer coefficients has the property that for all distinct primes p, q the number $f(p, q)$ is divisible by p or by q . Prove that $f(X, Y) = Xg(X, Y)$ or $f(X, Y) = Yg(X, Y)$ for some polynomial g with integer coefficients.

Proof. We need to prove that at least one of the polynomials $f(X, 0)$ and $f(0, Y)$ is 0. Assume that this is not the case and take positive integers c, d such that for all positive integers x

$$\max(|f(x, 0)|, |f(0, x)|) \leq cx^d.$$

This is possible, since $f(X, 0)$ and $f(0, X)$ are polynomials. Let S be the finite set of all roots of the polynomial $f(0, X)$ and consider a large positive integer N such that the equation $f(x, 0) = 0$ has no solution in $(cN^d, 2cN^d)$ (this holds for all sufficiently large N since by assumption $f(X, 0)$ is not the zero polynomial).

We claim that if $q \leq N$ and $p > cN^d$ are primes, then $q \in S$ or $q \mid f(p, 0)$. Indeed, suppose that $q \notin S$ and q does not divide $f(p, 0)$, thus q does not divide $f(p, q)$ and the hypothesis gives $p \mid f(p, q)$. This forces $p \mid f(0, q)$, which is impossible since $f(0, q) \neq 0$ and $|f(0, q)| \leq cq^d \leq cN^d < p$. The claim is therefore proved.

We conclude that for all primes $p > cN^d$

$$\prod_{q \leq N, q \notin S} q \mid f(p, 0).$$

By Bertrand's postulate there is a prime $p \in (cN^d, 2cN^d)$ and for such p the number $f(p, 0)$ is nonzero (by the choice of N) and $|f(p, 0)| \leq cp^d < c(2cN^d)^d$. We obtain therefore the existence of a constant k such that for all sufficiently large N we have

$$\prod_{q \leq N} q \leq kN^{d^2}.$$

This is however impossible by example 6.48. The result follows. \square

6.4 Problems for practice

The yoga of p -adic valuations

1. (Russia 2000) Prove that there is a partition of \mathbf{N} with 100 sets such that if $a, b, c \in \mathbf{N}$ satisfy $a + 99b = c$, then at least two of the numbers a, b, c belong to the same set.
2. (Iran 2012) Prove that for any positive integer t there is an integer $n > 1$ relatively prime to t such that none of the numbers $n+t, n^2+t, n^3+t, \dots$ is a perfect power.
3. Prove that if n, k are positive integers, then no matter how we choose signs \pm

$$\pm \frac{1}{k} \pm \frac{1}{k+1} \pm \dots \pm \frac{1}{k+n}$$

is not an integer.

4. (Romania TST 2007) Let $n \geq 3$ and let a_1, \dots, a_n be positive integers such that $\gcd(a_1, \dots, a_n) = 1$ and $\text{lcm}(a_1, \dots, a_n) \mid a_1 + \dots + a_n$. Prove that $a_1 a_2 \dots a_n$ divides $(a_1 + a_2 + \dots + a_n)^{n-2}$.
5. (Erdős-Turan) Let p be an odd prime and let S be a set of n positive integers. Prove that one can choose a subset T of S with at least $\lceil \frac{n}{2} \rceil$ elements such that for all distinct elements $a, b \in T$ we have

$$v_p(a+b) = \min(v_p(a), v_p(b)).$$

6. (Ostrowski) Find all functions $f : \mathbf{Q} \rightarrow [0, \infty)$ such that
- i) $f(x) = 0$ if and only if $x = 0$;
 - ii) $f(xy) = f(x) \cdot f(y)$ and $f(x+y) \leq \max(f(x), f(y))$ for all x, y .
7. Find all integers $n > 1$ for which

$$n^n \mid (n-1)^{n^{n+1}} + (n+1)^{n^{n-1}}.$$

8. (Mathlinks Contest) Let a, b be distinct positive rational numbers such that $a^n - b^n \in \mathbb{Z}$ for infinitely many positive integers n . Show that $a, b \in \mathbb{Z}$.
9. (Saint Petersburg) Find all positive integers m, n such that $m^n \mid n^m - 1$.
10. (Balkan 1993) Let p be a prime and let $m \geq 2$ be an integer. Prove that if the equation

$$\frac{x^p + y^p}{2} = \left(\frac{x+y}{2} \right)^m$$

has a positive integer solution $(x, y) \neq (1, 1)$, then $m = p$.

11. (China TST 2004) Let a be a positive integer. Prove that the equation $n! = a^b - a^c$ has a finite number of solutions (n, b, c) in positive integers.
12. (China TST 2016) Let c, d be integers greater than 1. Define a sequence $(a_n)_{n \geq 1}$ by $a_1 = c$ and $a_{n+1} = a_n^d + c$ for $n \geq 1$. Prove that for any $n \geq 2$ there is a prime number p dividing a_n and not dividing $a_1 a_2 \dots a_{n-1}$.

13. (Kvant M 1687) Find the largest possible number of elements of the set $\{2^n - 1 \mid n \in \mathbf{Z}\}$ that are terms of a geometric progression.
14. (Iran TST 2009) Let a be a positive integer. Prove that there are infinitely many primes dividing at least one of the numbers

$$2^{2^1} + a, 2^{2^2} + a, 2^{2^3} + a, \dots$$

15. (China TST 2016) A point in the coordinate plane is called rational if its coordinates are rational numbers. Given a positive integer n , can we color all rational points using n colors such that
- a) each point receives one color;
 - b) any line segment whose endpoints are rational points contains rational points of each of the n colors?
16. (China TST 2010) Let $k > 1$ be an integer and let $n = 2^{k+1}$. Prove that for any positive integers $a_1 < a_2 < \dots < a_n$, the number $\prod_{1 \leq i < j \leq n} (a_i + a_j)$ has at least $k + 1$ different prime divisors.

Legendre's formula

17. (Komal) Which binomial coefficients are powers of a prime?
18. Prove that $\binom{2n}{n} \mid \text{lcm}(1, 2, \dots, 2n)$ for all positive integers n .
19. Prove that for all positive integers n and all integers a we have

$$\frac{1}{n!} (a^n - 1)(a^n - a) \dots (a^n - a^{n-1}) \in \mathbf{Z}.$$

20. Prove that if $k < n$ then

$$n \binom{n-1}{k} \mid \text{lcm}(n, n-1, \dots, n-k).$$

21. (Mathematical Reflections S 206) Find all integers $n > 1$ having a prime factor p such that $v_p(n!) \mid n - 1$.

22. (Romania TST 2015) Let k be an integer greater than 1. When n runs through the integers greater than or equal to k , what is the largest number of divisors of $\binom{n}{k}$ that belong to $\{n - k + 1, n - k + 2, \dots, n\}$?
23. (Mathematical Reflections O 285) Define a sequence $(a_n)_{n \geq 1}$ by $a_1 = 1$ and $a_{n+1} = 2^n(2^{a_n} - 1)$ for $n \geq 1$. Prove that $n! \mid a_n$ for all $n \geq 1$.
24. (China 2015) For which integers k are there infinitely many positive integers n such that $n + k$ does not divide $\binom{2n}{n}$?
25. (Romania TST 2007) Find all positive integers x, y such that

$$x^{2007} - y^{2007} = x! - y!.$$

26. a) Prove that for all $n \geq 2$ we have

$$v_2 \left(\binom{4n}{2n} - (-1)^n \binom{2n}{n} \right) = s_2(n) + 2 + 3v_2(n),$$

where $s_2(n)$ is the sum of the digits in the base 2 expansion of n .

- b) (AMM E 2640) Find the exponent of 2 in the prime factorization of the number

$$\binom{2^{n+1}}{2^n} - \binom{2^n}{2^{n-1}}.$$

27. (China TST 2016) Define a function $f : \mathbf{N} \rightarrow \mathbf{Q}^*$ as follows: write a positive integer $n = 2^k m$ with $k \geq 0$ and m odd, and set $f(n) = m^{1-k}$. Prove that for all $n \geq 1$ the number $f(1)f(2)\dots f(n)$ is an integer divisible by any odd positive integer not exceeding n .
28. (IMO Shortlist 2014) If x is a real number, we denote by $\|x\|$ the distance between x and the nearest integer. Prove that if a, b are positive integers, then we can find a prime $p > 2$ and a positive integer k such that

$$\left\| \frac{a}{p^k} \right\| + \left\| \frac{b}{p^k} \right\| + \left\| \frac{a+b}{p^k} \right\| = 1.$$

29. (Erdős-Palfy-Szegedy theorem) Let a, b be positive integers such that the remainder of a when divided by any prime p does not exceed the remainder of b when divided by p . Prove that $a = b$.

Estimates for binomial coefficients and the distribution of prime numbers

30. Prove that there exist two consecutive squares such that there are at least 2000 primes between them.
31. A finite sequence of consecutive positive integers contains at least one prime number. Prove that the sequence contains a number that is relatively prime to all other terms of the sequence.
32. Prove that $2p_{n+1} \geq p_n + p_{n+2}$ for infinitely many n , where p_n is the n th prime.
33. (AMM) Find all integers $m, n > 1$ such that

$$1! \cdot 3! \cdot \dots \cdot (2n-1)! = m!.$$

34. (EMMO 2016) Let $a_1 < a_2 < \dots$ be an infinite increasing sequence of positive integers such that the sequence $\left(\frac{a_n}{n}\right)$ is bounded. Prove that for infinitely many n the number a_n divides $\text{lcm}(a_1, \dots, a_{n-1})$.
35. Does the equation $x! = y!(y+1)!$ have infinitely many solutions in positive integers?
36. (Richert's theorem) Prove that any integer larger than 6 is a sum of distinct primes.
37. (China TST 2015) Prove that there are infinitely many integers n such that $n^2 + 1$ is squarefree.
38. (USAMO 2014) Prove that there is a constant $c > 0$ with the following property: if a, b, n are positive integers such that $\gcd(a+i, b+j) > 1$ for all $i, j \in \{0, 1, \dots, n\}$, then

$$\min\{a, b\} > c^n \cdot n^{\frac{n}{2}}.$$

39. (Mertens) Prove that for all $n > 1$

$$-6 < \sum_{p \leq n} \frac{\ln p}{p} - \ln n < 4.$$

40. (Mertens) Prove that the sequence $(a_n)_{n \geq 2}$ defined by

$$a_n = \sum_{p \leq n} \frac{1}{p} - \ln \ln n$$

is bounded, where the sum is over all primes not exceeding n .

Chapter 7

Congruences for composite moduli

The goal of this chapter is to make a more detailed study of Euler's totient function and its applications to congruences for composite moduli. The first section deals with the Chinese remainder theorem, which we use to explain how to reduce polynomial congruences for composite moduli to congruences for primes and powers of primes (which was the subject of the previous chapter). We then establish Euler's theorem and give many applications. Finally, we discuss the important notion of order modulo n and that of primitive roots modulo n .

7.1 The Chinese remainder theorem

7.1.1 Proof of the theorem and first examples

The Chinese remainder theorem is a very useful result allowing one to find solutions to systems of linear congruences whose moduli are pairwise relatively prime. It is a very powerful tool in constructive problems. Roughly speaking, it says that congruences modulo a and modulo b are unrelated as long as a and b are relatively prime. The precise statement is the following.

Theorem 7.1. *Let m_1, m_2, \dots, m_k be pairwise relatively prime integers and let a_1, \dots, a_k be arbitrary integers. Then the system of congruences $x \equiv a_i \pmod{m_i}$, $1 \leq i \leq k$ has solutions, and these solutions form an infinite arithmetic progression with common difference $m_1 \dots m_k$ (in other words, any two solutions differ by a multiple of $m_1 \dots m_k$).*

Proof. For each $i \in \{1, 2, \dots, k\}$ we have $\gcd(m_i, \prod_{j \neq i} m_j) = 1$, thus there is an integer k_i such that $k_i \cdot \prod_{j \neq i} m_j \equiv 1 \pmod{m_i}$. Setting $x_i = k_i \cdot \prod_{j \neq i} m_j$, we have $x_i \equiv \delta_{ij} \pmod{m_j}$ for $1 \leq i, j \leq k$, where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. But then $x = a_1 x_1 + \dots + a_k x_k$ satisfies $x \equiv a_i \pmod{m_i}$ for $1 \leq i \leq k$, finishing the proof of the existence part.

Next, fix a solution x_0 of the system. Any other solution x satisfies $x \equiv a_i \pmod{m_i}$ for $1 \leq i \leq k$. Thus m_1, \dots, m_k divide $x - x_0$ and since they are relatively prime, we deduce that $m_1 \dots m_k \mid x - x_0$. Thus any two solutions differ by a multiple of $m_1 \dots m_k$. Conversely, if $m_1 \dots m_k \mid x - x_0$, then m_1, \dots, m_k all divide $x - x_0$ and so x is also a solution. Thus the solutions form an infinite arithmetic progression with common difference $m_1 \dots m_k$ and the result follows. \square

We continue with a long series of examples illustrating the Chinese remainder theorem. The condition that m_1, \dots, m_k are pairwise relatively prime may seem too strong in theorem 7.1. Note however that if $x, a_1, \dots, a_k, m_1, \dots, m_k$ are integers satisfying $x \equiv a_i \pmod{m_i}$ for $1 \leq i \leq k$, then necessarily $\gcd(m_i, m_j)$ divides $a_i - a_j = (x - a_j) - (x - a_i)$ for all $1 \leq i, j \leq k$. The next example states that this necessary condition is also sufficient, thereby establishing the optimal form of the Chinese remainder theorem.

Example 7.2. If a_1, a_2, \dots, a_k are integers and m_1, m_2, \dots, m_k are positive integers such that $a_i \equiv a_j \pmod{\gcd(m_i, m_j)}$ for all $1 \leq i, j \leq k$, then there are integers x such that $x \equiv a_i \pmod{m_i}$ for $1 \leq i \leq k$.

Proof. The result is clear if $m_1 m_2 \dots m_k = 1$, so assume that this is not the case and let p_1, \dots, p_n be the different prime factors of $m_1 m_2 \dots m_k$. For each i , choose $j(i)$ such that

$$v_{p_i}(m_{j(i)}) = \max(v_{p_i}(m_1), \dots, v_{p_i}(m_k))$$

and let $s_i = v_{p_i}(m_{j(i)})$.

By the Chinese remainder theorem we can find x such that $x \equiv a_{j(i)} \pmod{p_i^{s_i}}$ for all $1 \leq i \leq k$. We claim that x is a solution, which comes down to proving the inequality

$$v_{p_i}(x - a_l) \geq v_{p_i}(m_l)$$

for $1 \leq l \leq k$ and $1 \leq i \leq n$. By hypothesis $\gcd(m_l, m_{j(i)})$ divides $a_l - a_{j(i)}$, thus

$$v_{p_i}(a_l - a_{j(i)}) \geq v_{p_i}(\gcd(m_l, m_{j(i)})) = v_{p_i}(m_l).$$

It follows that

$$v_{p_i}(x - a_l) \geq \min(v_{p_i}(x - a_{j(i)}), v_{p_i}(a_{j(i)} - a_l)) \geq \min(s_i, v_{p_i}(m_l)) = v_{p_i}(m_l)$$

and we are done. \square

We continue with some constructive problems in which the Chinese remainder theorem plays a key role.

Example 7.3. (Czech-Slovak 2008) Prove the existence of a positive integer n such that for all integers k , all prime divisors of $k^2 + k + n$ are greater than 2008.

Proof. Let p_1, \dots, p_k be all primes not exceeding 2008. We deal first with each p_i , proving that we can find n such that the congruence $k^2 + k + n \equiv 0 \pmod{p_i}$ has no solutions. If $p_i = 2$, simply choose $n = 1$, so suppose that $p_i > 2$. Choose a quadratic non-residue a modulo p_i and pick n such that $4n - 1 \equiv -a \pmod{p_i}$ (which is possible since p_i is odd). Then the congruence $k^2 + k + n \equiv 0 \pmod{p_i}$ has no solutions, since any solution would satisfy $(2k + 1)^2 \equiv -(4n - 1) \equiv a \pmod{p_i}$, contradicting the choice of a . Thus we can find for each i an integer n_i such that the congruence $k^2 + k + n_i \equiv 0 \pmod{p_i}$ is not solvable. The Chinese remainder theorem shows that we can find n congruent to n_i modulo p_i for all $1 \leq i \leq k$, and such n satisfies all requirements by construction. \square

Example 7.4. (Russia 1995) Is there a permutation a_1, a_2, \dots of the set of all positive integers with the property that $a_1 + a_2 + \dots + a_n$ is a multiple of n for all $n \geq 1$?

Proof. We will prove that the answer is positive by inductively constructing such a sequence. Define $a_1 = 1$ and assume that a_1, \dots, a_k have already been defined. We will define next a_{k+1} and a_{k+2} . Let a_{k+2} be the smallest positive integer different from a_1, \dots, a_k . Next, choose a_{k+1} different from a_1, \dots, a_k, a_{k+2} such that $a_{k+1} \equiv -(a_1 + \dots + a_k) \pmod{k+1}$ and $a_{k+1} \equiv -(a_1 + \dots + a_k + a_{k+2}) \pmod{k+2}$. The existence of such a number is a consequence of the Chinese remainder theorem. Note that by construction the sequence a_1, a_2, \dots satisfies all requirements. \square

Example 7.5. (Baltic 2006) Is there a sequence a_1, a_2, a_3, \dots of positive integers such that the sum of every n consecutive elements is divisible by n^2 for every positive integer n ?

Proof. We will construct the sequence inductively. Set $a_1 = 1$ and suppose that a_1, \dots, a_k have already been constructed. For $1 \leq i \leq k$ let $b_i = (i+1)^2$ and $c_i = -a_k - a_{k-1} - \dots - a_{k-i+1}$. Note that if $i < j$, then $c_j - c_i$ is the sum of $j - i$ consecutive terms of the sequence a_1, \dots, a_k , hence a multiple of $(j - i)^2$, which itself is a multiple of $\gcd(b_i, b_j)$. By example 7.2 we can find a positive integer a_{k+1} such that $a_{k+1} \equiv c_i \pmod{b_i}$ for $1 \leq i \leq k$. Now any sequence of $j \in \{1, 2, \dots, k+1\}$ consecutive elements of the sequence a_1, \dots, a_{k+1} is a multiple of j^2 , so the inductive step is proved. \square

We end this section with some more challenging examples.

Example 7.6. (Russia 2008) Find all positive integers n with the following property: there are positive integers b_1, b_2, \dots, b_n , not all equal and such that the number $(b_1 + k)(b_2 + k) \dots (b_n + k)$ is a power of an integer for each natural number k . Here, a power means a number of the form x^y with $x, y > 1$.

Proof. If n is composite, say $n = ab$ with $a, b > 1$, then we can choose $b_1 = b_2 = \dots = b_a = 1$, then $b_{a+1} = \dots = b_{2a} = 2$ and all the other b_i 's equal to 1. Then for any k we have

$$(b_1 + k)(b_2 + k) \dots (b_n + k) = (k+1)^a (k+2)^a (k+1)^{a(b-2)},$$

which is a power.

Suppose now that n is a prime and that b_1, \dots, b_n satisfy the conditions of the problem. Let c_1, c_2, \dots, c_N be the set of distinct numbers among b_1, b_2, \dots, b_n , with multiplicities m_1, m_2, \dots, m_N . By assumption, we have $N > 1$ and clearly $n = m_1 + m_2 + \dots + m_N$. Moreover, for any k , the number $(c_1 + k)^{m_1}(c_2 + k)^{m_2} \dots (c_N + k)^{m_N}$ is a perfect power. The key point is to choose numbers k for which we can find distinct primes p_1, p_2, \dots, p_N such that $v_{p_i}(c_j + k) = 1$ if $i = j$ and 0 otherwise. In this case, if

$$(c_1 + k)^{m_1}(c_2 + k)^{m_2} \dots (c_N + k)^{m_N} = x^y$$

for some $x, y > 1$, we have $yv_{p_i}(x) = m_i$, so that y divides all m_i . But then y divides their sum, which is n and since n is a prime, it follows that $n = y$. Thus $n = y$ will divide all m_i and this obviously contradicts the fact that $N > 1$ and $m_1 + m_2 + \dots + m_N = n$.

Thus, we are done if we can find distinct primes p_1, p_2, \dots, p_N and k such that $v_{p_i}(c_j + k) = 1$ if $i = j$ and 0 otherwise. This is very simple: first, we choose some distinct prime numbers p_1, p_2, \dots, p_N , sufficiently large, say not dividing any of the numbers $c_i - c_j$ with $i \neq j$ and then choose k such that $k + c_i \equiv p_i \pmod{p_i^2}$ for all i . Such k exists by the Chinese remainder theorem. Of course, $v_{p_i}(k + c_i) = 1$ and for $j \neq i$ we cannot have $p_i | c_j + k$, since otherwise p_i would divide $c_i - c_j$, contradicting the choice of p_i . Thus, such k satisfies all desired conditions and the answer to the problem is: precisely the composite numbers. \square

Example 7.7. (IMO Shortlist 2014) Let $a_1 < a_2 < \dots < a_n$ be pairwise relatively prime positive integers with a_1 being prime and $a_1 \geq n + 2$. On the segment $I = [0, a_1 a_2 \dots a_n]$ of the real line, mark all integers that are divisible by at least one of the numbers a_1, a_2, \dots, a_n . These points split I into a number of smaller segments. Prove that the sum of the squares of the lengths of these segments is divisible by a_1 .

Proof. Let $0 = b_0 < b_1 < \dots < b_l = a_1 a_2 \dots a_n$ be all marked integers, thus we need to understand $(b_1 - b_0)^2 + (b_2 - b_1)^2 + \dots + (b_l - b_{l-1})^2$. We start by finding a more manageable expression.

Call an interval J admissible if it is a closed (nontrivial, i.e. not reduced to a point) sub-interval $[a, b]$ of $[0, a_1 \dots a_n]$ and there are no marked

points in the open interval (a, b) . Let N be the number of admissible intervals. Consider now pairs (I, J) , where I is an interval among $[b_0, b_1], [b_1, b_2], \dots, [b_{l-1}, b_l]$ and J is an admissible interval contained in I . Since the intervals $[b_0, b_1], [b_1, b_2], \dots, [b_{l-1}, b_l]$ have no common interior points and cover $[0, a_1 \dots a_n]$, for each admissible interval J there is a unique pair (I, J) attached to J , thus there are N such pairs. On the other hand, if we fix an interval I among $[b_0, b_1], [b_1, b_2], \dots, [b_{l-1}, b_l]$, say $I = [b_i, b_{i+1}]$ for some i , then clearly the admissible intervals contained in I are all intervals of the form $[x, y]$ with $b_i \leq x < y \leq b_{i+1}$, and there are $\binom{b_{i+1} - b_i + 1}{2}$ such intervals. Therefore, a double count of the pairs (I, J) reveals the crucial identity

$$\sum_{i=0}^{l-1} \binom{b_{i+1} - b_i + 1}{2} = N$$

or equivalently

$$\sum_{i=0}^{l-1} (b_{i+1} - b_i)^2 = 2N - a_1 \dots a_n.$$

It is therefore sufficient to prove that N is a multiple of a_1 . The advantage is that N is rather easily understood.

Since an admissible interval contains no multiple of a_1 in its interior, the length of the interval cannot exceed a_1 . Let us fix now $d \in \{1, 2, \dots, a_1\}$ and count the admissible intervals of length d . In other words, we need to find the number of integers $x \in \{0, 1, \dots, a_1 \dots a_n - d\}$ such that $(x, x + d)$ contains no multiple of any of the numbers a_1, \dots, a_n . Note that this is the same as the number of $x \in \{0, 1, \dots, a_1 \dots a_n - 1\}$ with the same property. Such x is a solution if and only if its remainder when divided by a_i belongs to $\{0, 1, \dots, a_i - d\}$ for all i . Since the numbers a_1, \dots, a_n are relatively prime, the Chinese remainder theorem implies that the number of such x is

$$f(d) = \prod_{i=1}^n (a_i - d + 1).$$

Thus

$$N = \sum_{d=1}^{a_1} (a_1 + 1 - d)(a_2 + 1 - d) \dots (a_n + 1 - d).$$

Since the polynomial $\prod_{i=1}^n (a_i + 1 - X)$ has degree $n < a_1 - 1$ and a_1 is a prime, corollary 5.77 yields

$$\sum_{d=1}^{a_1} (a_1 + 1 - d)(a_2 + 1 - d) \dots (a_n + 1 - d) \equiv 0 \pmod{a_1},$$

proving therefore that $a_1 \mid N$ and finishing the proof. \square

Example 7.8. (USA TST 2012) A function $f : \mathbf{N} \rightarrow \mathbf{N}$ has the property that $\gcd(f(m), f(n)) = 1$ whenever $\gcd(m, n) = 1$, and $n \leq f(n) \leq n + 2012$ for all n . Prove that if $n > 1$ then any prime divisor of $f(n)$ is a prime divisor of n .

Proof. We start by proving that f has many fixed points, more precisely we prove the existence of an infinite sequence $1 < j_1 < j_2 < \dots$ of pairwise relatively prime integers such that $f(j_k) = j_k$ for all k . Consider the sequence (a_n) defined by $a_1 = 2013! + 1$ and $a_{i+1} = a_i! + 1$ for $i \geq 1$. Then clearly a_1, a_2, \dots are pairwise relatively prime, so $f(a_1), f(a_2), \dots$ are also pairwise relatively prime. Since $0 \leq f(a_i) - a_i \leq 2012$ for all i , there is $k \in \{0, 1, \dots, 2012\}$ and an infinite sequence $i_1 < i_2 < \dots$ such that $f(a_{i_1}) - a_{i_1} = f(a_{i_2}) - a_{i_2} = \dots = k$. Since $k + 1 \mid a_i - 1 = a_{i-1}!$ for $i \geq 2$ (note that $a_j > 2013$ for all j), we have $k + 1 \mid a_{i_j} + k = f(a_{i_j})$ for all $j \geq 2$. Since $f(a_{i_2})$ and $f(a_{i_3})$ are relatively prime, this forces $k = 0$ and so we can take $j_1 = a_{i_1}, j_2 = a_{i_2}, \dots$, establishing the desired result. Note that since j_1, j_2, \dots are pairwise relatively prime, for any $N > 1$ there are infinitely many k such that $\gcd(j_k, N) = 1$.

Let now $n > 1$ and let p be a prime factor of $f(n)$. Suppose that p does not divide n . By the previous paragraph we can find pairwise relatively prime numbers $q_1 < \dots < q_{2012}$ which are relatively prime to $pn \cdot 2012!$ and satisfy $f(q_i) = q_i$ for $1 \leq i \leq 2012$. By the Chinese remainder theorem there is an integer $a > 1$ such that $a \equiv 0 \pmod{p}$, $a \equiv 1 \pmod{n}$ and $a \equiv -i \pmod{q_i}$ for $1 \leq i \leq 2012$. Since $\gcd(a, n) = 1$ and $p \mid \gcd(a, f(n))$, we cannot have $f(a) = a$, thus $f(a) = a + i$ for some $1 \leq i \leq 2012$. Then

$$\gcd(f(q_i), f(a)) = \gcd(q_i, a + i) > 1,$$

which gives $\gcd(q_i, a) > 1$. Combined with the congruence $a \equiv -i \pmod{q_i}$, this yields $\gcd(q_i, i) > 1$, which is impossible since $\gcd(q_i, 2012!) = 1$. Thus p must divide n and the result follows. \square

7.1.2 The local-global principle

The next theorem is very useful in practice: it shows that in order to solve polynomial congruences $f(x) \equiv 0 \pmod{n}$ it suffices to understand the case when n is a power of a prime, which we have already dealt with in chapter 4.

Theorem 7.9. *Let f be a polynomial with integer coefficients. If n is a positive integer, let*

$$A(n) = \{x \in \{0, 1, \dots, n-1\} \mid f(x) \equiv 0 \pmod{n}\}.$$

If m_1, \dots, m_k are pairwise relatively prime positive integers, then the map¹

$$A(m_1 \dots m_k) \rightarrow A(m_1) \times \dots \times A(m_k), \quad x \mapsto (x \pmod{m_1}, \dots, x \pmod{m_k})$$

is bijective. In particular, $A(m_1 \dots m_k)$ is nonempty if and only if $A(m_i)$ are nonempty for $1 \leq i \leq k$, in which case

$$|A(m_1 \dots m_k)| = |A(m_1)| \cdot \dots \cdot |A(m_k)|.$$

Proof. Let $n = m_1 \dots m_k$. Note that if $f(x) \equiv 0 \pmod{n}$ and $r_i \equiv x \pmod{m_i}$ then $0 \equiv f(x) \equiv f(r_i) \pmod{m_i}$, thus $r_i \in A(m_i)$ and the map, call it f , from the statement of the theorem is well-defined. Let us prove its injectivity. If $x, y \in A(n)$ have the same image through f then $x \equiv y \pmod{m_i}$ for $1 \leq i \leq k$. Since m_1, \dots, m_k are pairwise relatively prime, we deduce from the Chinese remainder theorem that x and y are congruent modulo $n = m_1 \dots m_k$. Since $x, y \in \{0, 1, \dots, n-1\}$ we conclude that $x = y$.

Let us prove now surjectivity. Let $x_i \in A(m_i)$, we need to prove the existence of $x \in A(n)$ such that $x \pmod{m_i} = x_i$ for $1 \leq i \leq k$. By the Chinese remainder theorem we can find $x \in \{0, 1, \dots, n-1\}$ such that $x \equiv x_i \pmod{m_i}$, thus $x \pmod{m_i} = x_i$ for $1 \leq i \leq k$. Since $x \equiv x_i \pmod{m_i}$ and $x_i \in A(m_i)$ we have $f(x) \equiv f(x_i) \equiv 0 \pmod{m_i}$ for $1 \leq i \leq k$. Using again that m_1, \dots, m_k are pairwise relatively prime, we deduce that $f(x) \equiv 0 \pmod{n}$ and so $x \in A(n)$, as desired. \square

The following result is an immediate consequence of the previous theorem, but we state it explicitly since it is very important in practice.

¹Here $x \pmod{N}$ denotes the remainder of x when divided by N

Corollary 7.10. *Let f be a polynomial with integer coefficients and let $n > 1$ be an integer, with prime factorization $n = p_1^{k_1} \dots p_s^{k_s}$. The number of solutions of the congruence $f(x) \equiv 0 \pmod{n}$ is the product of the number of solutions of the congruences $f(x) \equiv 0 \pmod{p_i^{k_i}}$, $1 \leq i \leq s$.*

Example 7.11. Let n be an integer greater than 1. Find the number of integers $x \in \{0, 1, \dots, n-1\}$ such that

- a) $x^2 \equiv x \pmod{n}$.
- b) $x^2 \equiv 1 \pmod{n}$.

Proof. a) We first consider the case when n is a power of a prime, say $n = p^k$ for some prime p and some $k \geq 1$. Then $x^2 \equiv x \pmod{n}$ is equivalent to $p^k \mid x(x-1)$. Since x and $x-1$ are relatively prime, this can only occur when either $p^k \mid x$ or $p^k \mid x-1$. In other words, in this case the congruence has exactly two solutions: 0 and 1. Corollary 7.10 then shows that in general the congruence $x^2 \equiv x \pmod{n}$ has 2^s solutions, where s is the number of distinct prime factors of n .

b) Similarly, we start with the case $n = p^k$, in which case we need to understand the divisibility $p^k \mid (x-1)(x+1)$. If $p > 2$ then p cannot divide both $x-1$ and $x+1$ thus we must have $p^k \mid x-1$ or $p^k \mid x+1$, giving two solutions ($x = 1$ and $x = p^k - 1$) of the congruence. Suppose now that $p = 2$. If $k = 1$ then we have one solution, $x = 1$, if $k = 2$ we have two solutions $x = 1$ and $x = 3$, so assume that $k \geq 3$. Then x must be odd and one of $x-1$, $x+1$ must be a multiple of 2^{k-1} since $\gcd(x-1, x+1) = 2$. We then obtain 4 solutions: $x = 1, 2^{k-1} + 1, 2^k - 1, 2^{k-1} - 1$. In conclusion, using corollary 7.10, we deduce that for $n = 2^\alpha p_1^{k_1} \dots p_s^{k_s}$ with p_1, \dots, p_s pairwise distinct odd primes and $k_i \geq 1$ (but we allow $s = 0$)

- if $\alpha \leq 1$ then the congruence has 2^s solutions.
- if $\alpha = 2$ the congruence has 2^{s+1} solutions.
- if $\alpha \geq 3$ the congruence has 2^{s+2} solutions. □

Example 7.12. Prove that the number of solutions of the congruence $x^2 \equiv -1 \pmod{n}$ is

- a) 0 if $4 \mid n$ or if $p \mid n$ for some prime $p \equiv 3 \pmod{4}$;
- b) 2^s otherwise, where s is the number of different odd prime divisors of n .

Proof. Part a) follows directly from corollary 5.28. For part b), by corollary 7.10 it suffices to deal with the case $n = p^k$ for some prime $p \equiv 1 \pmod{4}$ and some $k \geq 1$. In this case we need to prove that the congruence has exactly two solutions. The case $k = 1$ follows easily from theorem 5.55, and the general case follows from Hensel's lemma: each solution of the congruence $x^2 \equiv -1 \pmod{p}$ uniquely lifts to a solution of the congruence $x^2 \equiv -1 \pmod{p^k}$. \square

Example 7.13. Find all integers $n > 1$ for which we can find integers a, b such that

$$a^2 + b^2 + 1 \equiv 0 \pmod{n}.$$

Proof. Since $x^2 \equiv 0, 1 \pmod{4}$ for any integer x , the number $a^2 + b^2 + 1$ is never divisible by 4. Thus a solution n of the problem is not divisible by 4. Conversely, we will prove that if $n > 1$ is not a multiple of 4, then the congruence $a^2 + b^2 + 1 \equiv 0 \pmod{n}$ has solutions. Write $n = 2^e \cdot p_1^{e_1} \cdots p_s^{e_s}$ for some $e \in \{0, 1\}$, some pairwise distinct primes $2, p_1, \dots, p_s$ and some integers $e_1, \dots, e_s \geq 0$. If there are integers $a_0, b_0, \dots, a_s, b_s$ such that $a_0^2 + b_0^2 + 1 \equiv 0 \pmod{2^e}$ and $a_i^2 + b_i^2 + 1 \equiv 0 \pmod{p_i^{e_i}}$ for $1 \leq i \leq s$, then the Chinese remainder theorem gives us integers a, b such that $a \equiv a_0 \pmod{2^e}, a \equiv a_i \pmod{p_i^{e_i}}$ for $1 \leq i \leq s$ and similarly $b \equiv b_0 \pmod{2^e}, b \equiv b_i \pmod{p_i^{e_i}}$ for $1 \leq i \leq s$. Then clearly $a^2 + b^2 + 1 \equiv 0 \pmod{n}$. Thus we may assume that n is a power of a prime p , and $n \in \{1, 2\}$ if $p = 2$. The case $p = 2$ being clear, assume that $n = p^k$ with $p > 2$ and $k \geq 1$. We can find $a, b \in \{0, 1, \dots, \frac{p-1}{2}\}$ such that $a^2 \equiv -(b^2 + 1) \pmod{p}$ (since the sets $\{a^2 \pmod{p} | 0 \leq a \leq \frac{p-1}{2}\}$ and $\{(b^2 + 1) \pmod{p} | 0 \leq b \leq \frac{p-1}{2}\}$ have $\frac{p+1}{2}$ elements each, and there are $p < \frac{p+1}{2} + \frac{p+1}{2}$ remainders modulo p). Thus the congruence $a^2 + b^2 + 1 \equiv 0 \pmod{p}$ has solutions. Choose a solution (a_0, b_0) with $\gcd(p, a_0) = 1$ (we may always achieve this, possibly by permuting a_0 and b_0). Choose any integer b that is congruent to b_0 modulo p . Hensel's lemma applied to the polynomial $f(X) = X^2 + b^2 + 1$ shows that the solution a_0 modulo p of the congruence $f(x) \equiv 0 \pmod{p}$ lifts uniquely to a solution a modulo p^k of the congruence $f(x) \equiv 0 \pmod{p^k}$. Thus the congruence $a^2 + b^2 + 1 \equiv 0 \pmod{p^k}$ has solutions, and we are done. \square

Example 7.14. (generalization of IMO Shortlist 1997) Let $m, n > 1$ be relatively prime integers. An infinite arithmetic progression of integers contains an m th power and an n th power. Prove that it also contains an mn th power.

Proof. Let $(a + jd)_{j \geq 0}$ be the arithmetic progression. By assumption the congruences $x^m \equiv a \pmod{d}$ and $y^n \equiv a \pmod{d}$ have solutions, and we need to prove that the congruence $z^{mn} \equiv a \pmod{d}$ also has solutions. Using theorem 7.9, we may assume that $d = p^k$ for some prime p and some positive integer k . Choose integers x, y such that $x^m \equiv a \pmod{p^k}$ and $y^n \equiv a \pmod{p^k}$. If a is a multiple of p^k , simply take $z = 0$, so assume that $v_p(a) < k$. Since $x^m \equiv a \pmod{p^k}$, it follows that $mv_p(x) = v_p(x^m) = v_p(a)$. Similarly $nv_p(y) = v_p(a)$. Thus m and n divide $v_p(a)$ and hence mn also divides $v_p(a)$ (as m and n are relatively prime).

Write $v_p(a) = mnt$ for some integer t , thus $v_p(x) = nt$ and $v_p(y) = mt$. Since $x^m \equiv a \pmod{p^k}$, we deduce that $\frac{a}{p^{mnt}}$ is an m th power modulo p^{k-mnt} . Similarly, $\frac{a}{p^{mnt}}$ is an n th power modulo p^{k-mnt} . So, it suffices to prove the following lemma in order to conclude.

Lemma 7.15. *Let m, n be relatively prime, let p be a prime number and let $N \geq 1$. If x is relatively prime to p and is an m th power and an n th power modulo p^N , then it is also an mn th power modulo p^N .*

The proof of the lemma is very simple: choose integers a, b such that $x \equiv a^m \pmod{p^N}$ and $x \equiv b^n \pmod{p^N}$. Now $a^m \equiv b^n \pmod{p^N}$, hence $a^{um} \equiv b^{un} \pmod{p^N}$ for all $u \geq 1$. By Bezout's lemma we can find u such that $un \equiv 1 \pmod{m}$. The previous congruence shows that b must be an m th power modulo p^N , and so $x \equiv b^n \pmod{p^N}$ is an mn th power. \square

Example 7.16. Consider the polynomial $f(X) = (X^2 + 3)(X^2 - 13)(X^2 + 39)$. Prove that the congruence $f(x) \equiv 0 \pmod{n}$ has solutions for all integers $n > 1$.

Proof. By corollary 7.10 we may assume that n is a power of a prime, say $n = p^k$. Assume first that $k = 1$ and let us prove that at least one of the congruences $x^2 \equiv -3 \pmod{p}$, $x^2 \equiv 13 \pmod{p}$ and $x^2 \equiv -39 \pmod{p}$ has solutions. This is clear if $p = 3$ or $p = 13$, so assume that $\gcd(p, 39) = 1$. If

neither of these congruences has solutions, we obtain $\left(\frac{-3}{p}\right) = -1$, $\left(\frac{13}{p}\right) = -1$ and $\left(\frac{-39}{p}\right) = -1$, contradicting the multiplicative character of Legendre's symbol (theorem 5.101), which gives $\left(\frac{-39}{p}\right) = \left(\frac{-3}{p}\right) \cdot \left(\frac{13}{p}\right)$. This settles the case $k = 1$.

Assume now that $k > 1$ and $p \neq 2, 3, 13$. By Hensel's lemma any solution x_0 of the congruence $x^2 \equiv a \pmod{p}$ with $a \in \{-3, 13, -39\}$ lifts uniquely to a solution of the congruence $x^2 \equiv a \pmod{p^k}$ (note that $2x_0$ is not divisible by p by our hypothesis on p). Thus we are done in this case. It remains to deal with the cases $p = 2, 3, 13$. If $p = 3$ we can use Hensel's lemma to lift the solution $x = 1$ of the congruence $x^2 \equiv 13 \pmod{3}$ to a solution of the congruence $x^2 \equiv 13 \pmod{3^k}$. We deal similarly with the case $p = 13$, by lifting via Hensel's lemma the solution $x = 6$ of the congruence $x^2 + 3 \equiv 0 \pmod{13}$. Finally, we have to deal with the case $p = 2$. We prove by induction the existence of a sequence x_n such that $2^n | x_n^2 + 39$. Take $x_1 = 1$, $x_2 = 1$ and $x_3 = 1$. Assuming that $x_n^2 + 39 = 2^n \cdot k$ for some integer k and $n \geq 3$, we have $(2^{n-1}x + x_n)^2 + 39 = 2^n(xx_n + k) \pmod{2^{n+1}}$. If k is even set $x_{n+1} = x_n$, otherwise set $x = 1$ and so $x_{n+1} = x_n + k$. Note that the case $p = 2$ could also have been treated using example 5.170(b), where we saw that the congruence $x^2 \equiv a \pmod{2^n}$ has solutions for all n if $a \equiv 1 \pmod{8}$. Applying this to $a = -39$ solves this case. \square

The next example is a variation on the proof of theorem 7.9.

Example 7.17. (AMM E 2330) Let $f : \mathbf{N} \rightarrow \mathbf{Z}$ be a function such that $a - b \mid f(a) - f(b)$ for all positive integers a, b . Let $a(n)$ (respectively $b(n)$) be the number of terms of the sequence $f(1), f(2), \dots, f(n)$ which are multiples of n (respectively relatively prime to n). Prove that $a, b : \mathbf{N} \rightarrow \mathbf{Z}$ are multiplicative functions and

$$b(n) = n \prod_{p|n} \left(1 - \frac{a(p)}{p}\right).$$

Proof. We start with a simple but crucial observation. Let m, n be relatively prime integers and consider $j \in \{1, 2, \dots, mn\}$. Let $u \in \{1, 2, \dots, n\}$ and $v \in \{1, 2, \dots, m\}$ be the unique integers for which $j \equiv u \pmod{n}$ and $j \equiv v \pmod{m}$. Then mn divides $f(j)$ if and only if $n \mid f(u)$ and $m \mid f(v)$.

Indeed, mn divides $f(j)$ if and only if $m \mid f(j)$ and $n \mid f(j)$, which happens if and only if $m \mid f(v)$ and $n \mid f(u)$ (since $f(j) \equiv f(v) \pmod{m}$ and $f(j) \equiv f(u) \pmod{n}$ by assumption).

Next, let

$$A = \{u \in \{1, 2, \dots, n\} \mid n \mid f(u)\}, \quad B = \{v \in \{1, 2, \dots, m\} \mid m \mid f(v)\}.$$

For each $(u, v) \in A \times B$ there is a unique integer $j(u, v) \in \{1, 2, \dots, mn\}$ such that $j(u, v) \equiv u \pmod{n}$ and $j(u, v) \equiv v \pmod{m}$, by the Chinese remainder theorem. By the previous paragraph, the numbers $j(u, v)$ with (u, v) running through $A \times B$ are exactly the integers $j \in \{1, 2, \dots, mn\}$ such that $mn \mid f(j)$, yielding $a(mn) = a(m)a(n)$.

Next, for each prime divisor p of n let A_p be the set of numbers $j \in \{1, 2, \dots, n\}$ such that $p \mid f(j)$. The inclusion-exclusion principle yields

$$b(n) = n - |\cup_{p|n} A_p| = n - \sum_{p|n} |A_p| + \sum_{p \neq q|n} |A_{pq}| + \dots$$

If d is a positive divisor of n , then for each $s \geq 0$ there are $a(d)$ integers k between $sd + 1$ and $(s + 1)d$ for which $d \mid f(k)$ (this follows from the definition of $a(d)$ and the fact that $d \mid f(j)$ if and only if $d \mid f(u_j)$, where $u_j \in \{1, 2, \dots, d\}$ is the unique integer congruent to j modulo d). Thus there are $\frac{n}{d}a(d)$ integers $j \in \{1, 2, \dots, n\}$ such that $d \mid f(j)$. Thus if p_1, \dots, p_s are pairwise distinct prime divisors of n then

$$|A_{p_1 \dots p_s}| = \frac{n}{p_1 \dots p_s} a(p_1 \dots p_s) = \frac{n}{p_1 \dots p_s} a(p_1) \dots a(p_s)$$

which combined with the previous formula for $b(n)$ yields

$$b(n) = n \prod_{p|n} \left(1 - \frac{a(p)}{p}\right).$$

It is clear from this last formula that $n \mapsto b(n)$ is multiplicative. □

The result established in the previous example is fairly useful, as the following two examples show.

Example 7.18. Prove that for any integer $n > 1$ the number of integers $a \in \{1, 2, \dots, n\}$ such that a and $a+1$ are both relatively prime to n is $n \prod_{p|n} \left(1 - \frac{2}{p}\right)$.

Proof. Take $f(x) = x(x+1)$ and apply example 7.17. For each prime p there are exactly 2 integers $k \in \{1, 2, \dots, p\}$ such that $p \mid f(k)$, namely $k = p-1$ and $k = p$, thus with the notations of example 7.17 we have $a(p) = 2$ for all primes p . The result follows. \square

Example 7.19. (Menon's identity) Prove that for any integer $n > 1$

$$\sum_{\substack{1 \leq k \leq n \\ (k,n)=1}} \gcd(n, k-1) = \phi(n)\tau(n).$$

Proof. Using Gauss' theorem 4.112 we obtain

$$\sum_{\substack{1 \leq k \leq n \\ (k,n)=1}} \gcd(n, k-1) = \sum_{\substack{1 \leq k \leq n \\ (k,n)=1}} \sum_{e|\gcd(n, k-1)} \varphi(e) = \sum_{e|n} \varphi(e) \sum_{k \in S(e)} 1 = \sum_{e|n} \varphi(e) |S(e)|,$$

where $S(e)$ is the set of integers $k \in \{1, \dots, n\}$ which are relatively prime to n and satisfy $k \equiv 1 \pmod{e}$. It suffices to prove that $S(e)$ has $\frac{\varphi(n)}{\varphi(e)}$ elements for all $e \mid n$. Fix such e and note that $S(e)$ is in bijection with the set of $x \in \{0, 1, \dots, \frac{n}{e} - 1\}$ for which $1 + xe$ is relatively prime to n (simply set $k = 1 + xe$) or equivalently $1 + xe$ is relatively prime to $\frac{n}{e}$. Applying example 7.17 to $f(x) = 1 + xe$ with $\frac{n}{e}$ instead of n and noting that the number of multiples of p among $f(1), \dots, f(\frac{n}{e})$ is 1 when p does not divide e and 0 otherwise, we obtain

$$|S(e)| = \frac{n}{e} \prod_{\substack{p|\frac{n}{e} \\ \gcd(p,e)=1}} \left(1 - \frac{1}{p}\right) = \frac{n}{e} \cdot \frac{\prod_{p|n} \left(1 - \frac{1}{p}\right)}{\prod_{p|e} \left(1 - \frac{1}{p}\right)} = \frac{\varphi(n)}{\varphi(e)},$$

as desired. \square

We end this section with a more difficult result, which is also quite useful in practice.

Theorem 7.20. *If an integer a is a quadratic residue modulo all sufficiently large primes, then a is a perfect square.*

Proof. First note that if p^2 divides a for some prime p , then a/p^2 will also be a quadratic residue modulo all sufficiently large primes. Thus we may assume a has no repeated prime factors and hence $a = \pm p_1 p_2 \dots p_s$ for pairwise distinct primes p_1, \dots, p_s .

Suppose p_s is odd. Let r be a quadratic non-residue modulo p_s . By the Chinese remainder theorem the solutions of the simultaneous congruences

$$q \equiv 1 \pmod{8p_1 \dots p_{s-1}}, \quad \text{and} \quad q \equiv r \pmod{p_s}$$

form an infinite arithmetic progression $x + 8p_1 \dots p_s \mathbf{Z}$ for some integer x . Clearly $\gcd(x, 8p_1 \dots p_s) = 1$ and so by Dirichlet's theorem this arithmetic progression contains infinitely many primes q . Since such a prime q is 1 modulo 8 by construction, we have $\left(\frac{\pm 2}{q}\right) = 1$. Also $(-1)^{(q-1)/2} = 1$, so the quadratic reciprocity law gives $\left(\frac{p_i}{q}\right) = \left(\frac{q}{p_i}\right)$, which equals 1 if $i \neq s$ and -1 for $i = s$. Thus

$$\left(\frac{a}{q}\right) = \left(\frac{\pm 1}{q}\right) \prod_{i=1}^s \left(\frac{p_i}{q}\right) = -1,$$

contradicting the choice of a .

Thus a has no odd prime factors and hence $a = \pm 1$ or ± 2 . However if q is a large prime congruent to 3 modulo 8, then $\left(\frac{-1}{q}\right) = \left(\frac{2}{q}\right) = -1$, and if q is 5 modulo 8, then $\left(\frac{-2}{q}\right) = -1$. Thus the only possibility is $a = 1$. Since we only cancelled off squares of primes, it follows that our original a was a perfect square. \square

Example 7.21. A quadratic polynomial f with integer coefficients has the property that for any prime p the congruence $f(n) \equiv 0 \pmod{p}$ has at least one solution. Prove that f has a rational root.

Proof. Writing $f(X) = aX^2 + bX + c$, we need to prove that $\Delta := b^2 - 4ac$ is a perfect square. Let p be any prime and let n be an integer such that $f(n) \equiv 0 \pmod{p}$, then

$$\Delta \equiv 4af(n) + \Delta = (2an + b)^2 \pmod{p}$$

and so Δ is a quadratic residue modulo p . The result follows then from theorem 7.20. \square

Example 7.22. (Mathlinks Contest) Nonnegative integers $a_1, a_2, \dots, a_{2004}$ have the property that $a_1^n + a_2^n + \dots + a_{2004}^n$ is a perfect square for all positive integers n . What is the least number of terms of the sequence $a_1, a_2, \dots, a_{2004}$ that are equal to 0?

Proof. Suppose that b_1, \dots, b_k are positive integers such that $b_1^n + b_2^n + \dots + b_k^n$ is a perfect square for all n . If p is a prime not dividing $b_1 b_2 \dots b_k$, then Fermat's little theorem gives

$$b_1^{p-1} + b_2^{p-1} + \dots + b_k^{p-1} \equiv k \pmod{p}$$

and the left-hand side is a perfect square, thus k is a quadratic residue modulo p . It follows from theorem 7.20 that k is a perfect square. Since the greatest perfect square smaller than 2004 is $44^2 = 1936$, there must be at least $2004 - 1936 = 68$ zeros in the sequence a_1, \dots, a_{2004} . To see that this is optimal, simply take $a_1 = \dots = a_{1936} = 1$ and the other terms equal to 0. \square

7.1.3 Covering systems of congruences

We discuss in this section a topic closely related to the Chinese remainder theorem, that of covering systems of congruences. These were introduced by Erdős in order to give an explicit construction of an infinite arithmetic progression of positive integers none of whose terms can be written in the form $2^k + p$ with $k \geq 0$ and p a prime number. This problem has a quite long history: de Polignac conjectured in 1849 that any odd integer $n > 1$ can be written $n = 2^k + p$ with $k \geq 0$ and p either a prime number or equal to 1. This conjecture turns out to be false, for instance 127 and 905 are counterexamples. Using covering systems of congruences and a very clever application of the Chinese remainder theorem, Erdős constructed an explicit infinite arithmetic progression all of whose terms are counterexamples to de Polignac's conjecture (it was known previously, thanks to work of van der Corput, that a positive proportion of the odd integers are counterexamples).

We will discuss his construction in this section, as well as some other results related to covering systems of congruences.

If a and n are integers with $n > 1$, we write

$$a + n\mathbf{Z} = \{a + nx \mid x \in \mathbf{Z}\}$$

for the infinite arithmetic progression consisting of numbers congruent to a modulo n . In other words $a + n\mathbf{Z}$ is the residue class of a modulo n .

Definition 7.23. A covering system is a finite collection of arithmetic progressions $a_1 + n_1\mathbf{Z}, \dots, a_k + n_k\mathbf{Z}$, with $a_1, \dots, a_k \in \mathbf{Z}$ and $n_1, \dots, n_k > 1$, such that

$$\mathbf{Z} = \bigcup_{i=1}^k (a_i + n_i\mathbf{Z}).$$

The numbers n_1, \dots, n_k are called the moduli of the covering system (note that we impose the condition $n_1, \dots, n_k > 1$ to avoid trivial considerations in the sequel).

A trivial covering system of congruences is obtained as follows: choose any $N > 1$ and consider the arithmetic progressions $(i + N\mathbf{Z})_{1 \leq i \leq N}$. This is certainly not very impressive, so let us give a few other examples:

a) An interesting covering system with distinct moduli (and smallest modulus 2) is

$$2\mathbf{Z}, 3\mathbf{Z}, 1 + 4\mathbf{Z}, 5 + 6\mathbf{Z}, 7 + 12\mathbf{Z}.$$

The reader will easily convince himself that this is indeed a covering system.

b) Erdős' construction (to be given below) uses the covering system given by

$$2\mathbf{Z}, 3\mathbf{Z}, 1 + 4\mathbf{Z}, 3 + 8\mathbf{Z}, 7 + 12\mathbf{Z}, 23 + 24\mathbf{Z}.$$

It is not difficult, although a bit tedious, to check that this is indeed a covering system.

c) A covering system, due to Davenport and Erdős, with smallest modulus 3 and distinct moduli is given by

$$\begin{aligned} &3\mathbf{Z}, 4\mathbf{Z}, 5\mathbf{Z}, 1 + 6\mathbf{Z}, 6 + 8\mathbf{Z}, 3 + 10\mathbf{Z}, 5 + 12\mathbf{Z}, 11 + 15\mathbf{Z}, \\ &7 + 20\mathbf{Z}, 10 + 24\mathbf{Z}, 2 + 30\mathbf{Z}, 34 + 40\mathbf{Z}, 59 + 60\mathbf{Z}, 98 + 120\mathbf{Z}. \end{aligned}$$

d) Here is yet another example, due to Erdős:

$$2\mathbf{Z}, 3\mathbf{Z}, 5\mathbf{Z}, 1 + 6\mathbf{Z}, 7\mathbf{Z}, 1 + 10\mathbf{Z}, 1 + 14\mathbf{Z}, 2 + 15\mathbf{Z}, 2 + 21\mathbf{Z}, \\ 23 + 30\mathbf{Z}, 4 + 35\mathbf{Z}, 5 + 42\mathbf{Z}, 59 + 70\mathbf{Z}, 104 + 105\mathbf{Z}.$$

As the reader has already guessed, it takes a bit more work to check that these last two examples are indeed covering systems.

Probably influenced by the previous examples, Erdős conjectured that for any N one can find a covering system of congruences with distinct moduli and in which the smallest modulus is greater than N . Choi constructed in 1971 a covering whose smallest modulus is 20, and one had to wait until 2006 for the construction (by Gibson) of a covering system with smallest modulus 25. In 2009 Nielsen proved the existence of a covering system with smallest modulus 40. All this suggested that Erdős' conjecture is true. In a spectacular work, Bob Hough proved in 2015 the following result, which disproves Erdős' conjecture.

Theorem 7.24. *(Bob Hough) In every covering system with distinct moduli, the smallest modulus cannot exceed 10^8 .*

There are many open problems concerning covering systems, some of which look surprisingly innocent. For instance, the Erdős-Selfridge conjecture states that there is no covering system whose moduli are distinct odd integers (greater than 1).

We are now ready to present Erdős' clever argument.

Theorem 7.25. *(Erdős) There is an infinite arithmetic progression consisting of odd positive integers n which cannot be written as the sum of a power of 2 and of a prime number.*

Proof. We will use the covering system

$$2\mathbf{Z}, 3\mathbf{Z}, 1 + 4\mathbf{Z}, 3 + 8\mathbf{Z}, 7 + 12\mathbf{Z}, 23 + 24\mathbf{Z},$$

which we represent as $(a_i + n_i\mathbf{Z})_{1 \leq i \leq k}$ (so $a_1 = 0$, $a_2 = 0$, $a_3 = 1$, $n_1 = 2$, $n_2 = 3$, $n_3 = 4$, etc). Next, choose pairwise distinct primes p_1, \dots, p_k such that $p_i \mid 2^{n_i} - 1$ for all i . This is possible, for instance by using the fact that

$$3 \mid 2^2 - 1, 7 \mid 2^3 - 1, 5 \mid 2^4 - 1, 17 \mid 2^8 - 1, 13 \mid 2^{12} - 1, 241 \mid 2^{24} - 1$$

we can choose

$$p_1 = 3, p_2 = 7, p_3 = 5, p_4 = 17, p_5 = 13, p_6 = 241.$$

Using the Chinese remainder theorem, we can find an infinite arithmetic progression of (odd and positive) integers n such that

$$n \equiv 1 \pmod{2^{241}}, \quad n \equiv 2^{a_i} \pmod{p_i}, \quad 1 \leq i \leq 6.$$

We claim that any such integer n which is greater than $2^{241} + 241$ is not of the form $2^k + p$ with $k \geq 0$ and p a prime number. Indeed, suppose that $n = 2^k + p$ and choose i such that $k \equiv a_i \pmod{n_i}$. Then $2^k \equiv 2^{a_i} \pmod{2^{n_i} - 1}$, thus $2^k \equiv 2^{a_i} \pmod{p_i}$. Since $n \equiv 2^{a_i} \pmod{p_i}$, we deduce that $p \equiv 0 \pmod{p_i}$ and so necessarily $p = p_i$. Since $n > 2^{241} + 241$ and $p_i \leq 241$, we have $k > 241$. But then taking the equation $n = 2^k + p_i$ modulo 2^{241} yields $1 \equiv p_i \pmod{2^{241}}$, which is certainly impossible since $p_i \leq 241$. \square

The next example uses a very similar argument.

Example 7.26. (Sierpinski-Selfridge) Prove that there is a positive integer k such that $k \cdot 2^n + 1$ is composite for all positive integers n .

Proof. Let $F_n = 2^{2^n} + 1$ be the n th Fermat's number. Write $F_5 = ab$ with $a, b > 1$ (one can take $a = 641$, see example 2.12). Since the Fermat numbers are pairwise relatively prime, (example 3.12), the Chinese remainder theorem yields infinitely many positive integers k such that

$$k \equiv 2 \pmod{F_0 F_1 F_2 F_3 F_4 a}, \quad \text{and} \quad k \equiv -2 \pmod{b}.$$

We will prove that for each $n \geq 0$ one of the numbers a, b, F_0, \dots, F_4 divides $k \cdot 2^n + 1$. Let $j = v_2(n + 1)$ and write $n = s \cdot 2^j - 1$ for an odd number s . We will discuss three cases. If $j > 5$, then $k \cdot 2^n + 1 \equiv -2^{n+1} + 1 \pmod{b}$ and b divides F_5 , which divides $2^{2^5} - 1$, which finally divides $2^{n+1} - 1$, hence $b \mid k \cdot 2^n + 1$. If $j = 5$, then since $a \mid F_5$, we have

$$k \cdot 2^n + 1 \equiv 2^{n+1} + 1 = 2^{2^5 s} + 1 \equiv 0 \pmod{a}.$$

Similarly, if $j \leq 4$, then

$$k \cdot 2^n + 1 \equiv 2^{n+1} + 1 = 22^j s + 1 \equiv 0 \pmod{F_j}.$$

We are now done: simply choose $k > F_5$ satisfying the previous congruences. Then for all $n \geq 0$ the number $k \cdot 2^n + 1$ is greater than each of the numbers F_0, \dots, F_4, a, b , and divisible by at least one of them. Hence $k \cdot 2^n + 1$ is composite for all n and we are done. \square

Remark 7.27. a) The result established in the previous example was obtained by Sierpinski in 1960. His approach (which is the one explained above) gave an infinite family of solutions, namely all

$$k \equiv 15511380746462593381 \pmod{2 \cdot 3 \cdot 5 \cdot 17 \cdot 257 \cdot 65537 \cdot 641 \cdot 6700417}.$$

In 1962 Selfridge found that $78557 \cdot 2^n + 1$ is composite for all $n \geq 1$, being always a multiple of one of the numbers 3, 5, 7, 13, 19, 37 or 73. This is based on the fact that

$$2\mathbf{Z}, 1 + 4\mathbf{Z}, 3 + 9\mathbf{Z}, 15 + 18\mathbf{Z}, 27 + 36\mathbf{Z}, 1 + 3\mathbf{Z}, 11 + 12\mathbf{Z}$$

is a covering system, and on the fact that $x = 78557$ is a solution of the following congruences

$$x \equiv 2 \pmod{3}, x \equiv 2 \pmod{5}, x \equiv 9 \pmod{73}, x \equiv 11 \pmod{19},$$

$$x \equiv 6 \pmod{37}, x \equiv 3 \pmod{7}, x \equiv 11 \pmod{13}.$$

For instance, if $n \in 2\mathbf{Z}$ then $x \cdot 2^n + 1$ is a multiple of 3, if $n \in 1 + 4\mathbf{Z}$ then $x \cdot 2^n + 1$ is a multiple of 5, ..., if $n \in 11 + 12\mathbf{Z}$ then $x \cdot 2^n + 1$ is a multiple of 13. Conjecturally, 78557 is the smallest positive integer k for which $k \cdot 2^n + 1$ is composite for all n (it is known that there can be at most five possible smaller numbers).

b) We could also have proved this result using the covering system from the proof of Erdős' theorem and a similar argument. Reversing the signs of the congruences would yield infinitely many n such that for all k the number

$n + 2^k$ is divisible by one of the primes 3, 5, 7, 13, 17, 241. But then for any such n and any k the number

$$n + 2^{k((3-1)(5-1)(7-1)(13-1)(17-1)(241-1)-1)}$$

is divisible by some prime $p \in \{3, 5, 7, 13, 17, 241\}$, and Fermat's little theorem yields $p \mid n \cdot 2^k + 1$.

Example 7.28. Let $(a_i + n_i\mathbf{Z})$ be a covering system with pairwise distinct moduli $n_1, \dots, n_k > 1$. Prove that the arithmetic progressions $a_1 + n_1\mathbf{Z}$, $a_2 + n_2\mathbf{Z}$, \dots , $a_k + n_k\mathbf{Z}$ are not pairwise disjoint.

Proof. Assume that the progressions are pairwise disjoint and let

$$N = \text{lcm}(n_1, \dots, n_k) \text{ and } \zeta_N = e^{\frac{2i\pi}{N}}.$$

For each $1 \leq j \leq k$ let

$$z_j = e^{\frac{2i\pi a_j}{n_j}}.$$

It is not difficult to check that the solutions of the equation

$$x^{\frac{N}{n_j}} = z_j$$

are precisely the numbers ζ_N^u with $u \in a_j + n_j\mathbf{Z}$. Since the arithmetic progressions $a_1 + n_1\mathbf{Z}$, $a_2 + n_2\mathbf{Z}$, \dots , $a_k + n_k\mathbf{Z}$ are pairwise disjoint and their union is \mathbf{Z} , we deduce that

$$X^N - 1 = P_1 P_2 \dots P_k, \quad \text{where} \quad P_j(X) = X^{\frac{N}{n_j}} - z_j.$$

Indeed, it follows from the above description of the roots of P_1, \dots, P_k that $X^N - 1$ and $P_1 \dots P_k$ have exactly the same roots, with the same multiplicity, namely 1.

By symmetry, we may assume that $n_k > \dots > n_1$, so that $\frac{N}{n_1} > \dots > \frac{N}{n_k}$. The coefficient of $X^{\frac{N}{n_k}}$ in the right-hand side of the equality

$$X^N - 1 = (X^{\frac{N}{n_1}} - z_1) \cdot \dots \cdot (X^{\frac{N}{n_k}} - z_k)$$

is $(-1)^{k-1}z_1 \dots z_{k-1}$, while the coefficient of $X^{\frac{N}{n_k}}$ in the left-hand side is 0. We obtain $(-1)^{k-1}z_1 \dots z_{k-1} = 0$, which is obviously impossible. The result follows. \square

The reader will compare the next result with the one established in example 7.2.

Example 7.29. (AMM 5747) Let $1 < n_1 < \dots < n_k$ be integers and let $0 \leq b_i < n_i$ be integers for all $1 \leq i \leq k$. Assuming that $\gcd(n_i, n_j)$ does not divide $b_i - b_j$ for all $i \neq j$, prove the existence of an integer x which is not congruent to b_i modulo n_i for all $1 \leq i \leq k$.

Proof. Assume that this is not the case, so any integer x satisfies one of the congruences $x \equiv b_i \pmod{n_i}$, in other words $(b_i + n_i \mathbf{Z})_{1 \leq i \leq k}$ define a covering system. Note that if $i \neq j$, then x cannot satisfy simultaneously $x \equiv b_i \pmod{n_i}$ and $x \equiv b_j \pmod{n_j}$, for otherwise we would obtain $\gcd(n_i, n_j) \mid b_i - b_j$, contradicting the hypothesis. The result follows then immediately from the previous example. \square

Example 7.30. (Erdős-Sun) A family of k arithmetic progressions

$$(a_i + n_i \mathbf{Z})_{1 \leq i \leq k}$$

(with a_i, n_i integers and $n_i > 1$) has the property that $\bigcup_{i=1}^k (a_i + n_i \mathbf{Z})$ contains 2^k consecutive integers. Prove that this family is a covering system of congruences.

Proof. The key observation is that an integer x belongs to $\bigcup_{i=1}^k (a_i + n_i \mathbf{Z})$ if and only if

$$\prod_{j=1}^k \left(1 - e^{\frac{2i\pi}{n_j}(x-a_j)} \right) = 0.$$

A brutal expansion of the left-hand side yields

$$\prod_{j=1}^k \left(1 - e^{\frac{2i\pi}{n_j}(x-a_j)} \right) = \sum_{I \subset \{1, 2, \dots, k\}} c_I \cdot e^{2i\pi x d_I},$$

where

$$c_I = (-1)^{|I|} \prod_{j \in I} e^{-2i\pi \frac{a_j}{n_j}}, \quad d_I = \sum_{j \in I} \frac{1}{n_j}.$$

the sum being taken over all subsets I of $\{1, 2, \dots, k\}$ (with the convention that the product over the empty set is 1). Note that c_I, d_I are complex numbers depending only on the family of arithmetic progressions and not on x . Letting $z_I = e^{2i\pi d_I}$, the hypothesis says that the relation $\sum_I c_I z_I^x = 0$ holds for 2^k consecutive integers x , and we need to prove that it holds for all integers x . Letting

$$u_n = \sum_I c_I z_I^n,$$

it follows that 2^k consecutive terms of this sequence vanish. On the other hand, the sequence $(u_n)_n$ satisfies a linear recurrence relation with constant coefficients, of order 2^k . Indeed, writing

$$\prod_I (X - z_I) = X^{2^k} + A_{2^k-1} X^{2^k-1} + \dots + A_1 X + A_0,$$

we have the recurrence relation

$$u_{n+2^k} + A_{2^k-1} u_{n+2^k-1} + \dots + A_0 u_n = 0.$$

Since $A_0 \neq 0$ and since by assumption 2^k consecutive terms of this sequence vanish, it follows immediately that all terms vanish, which is what we needed. \square

Example 7.31. (Zhang's theorem) Prove that for any covering system of congruences $(a_i + n_i \mathbf{Z})_{1 \leq i \leq k}$ there exists a nonempty subset $I \subseteq \{1, \dots, k\}$ such that

$$\sum_{i \in I} \frac{1}{n_i} \in \mathbf{Z}$$

Proof. An argument identical to the one used in the proof of example 7.30 yields for all integers n

$$1 + \sum_{I \subset \{1, 2, \dots, k\}} c_I \cdot e^{2i\pi n d_I} = 0,$$

where now the sum is over the nonempty subsets I of $\{1, 2, \dots, k\}$ and

$$c_I = (-1)^{|I|} \prod_{j \in I} e^{-2i\pi \frac{a_j}{n_j}}, \quad d_I = \sum_{j \in I} \frac{1}{n_j}.$$

We need to prove that at least one of the numbers d_I is an integer. The key observation is the following

Lemma 7.32. *Suppose that $x \in \mathbf{R}$ is not an integer. Then the sequence $(a_n)_{n \geq 1}$ defined by*

$$a_n = \sum_{k=1}^n e^{2i\pi x k}$$

is bounded.

Proof. Write $z = e^{2i\pi x}$ and observe that $z \neq 1$ since x is not an integer. Then

$$a_n = z + z^2 + \dots + z^n = z \cdot \frac{1 - z^n}{1 - z}$$

and since $|z| = 1$ it is clear that

$$|a_n| \leq \frac{2}{|1 - z|}.$$

The result follows. □

Assuming next that none of the numbers d_I is an integer, we obtain a contradiction using the lemma and the following relation, which follows by adding the previous ones for $n = 1, 2, \dots, N$:

$$-N = \sum_{I \subset \{1, 2, \dots, k\}} c_I \cdot \sum_{n=1}^N e^{2i\pi n d_I}.$$

Indeed, the left-hand side is obviously unbounded as $N \rightarrow \infty$, while the right-hand side is bounded thanks to the lemma and our assumption. The result follows. □

7.2 Euler's theorem

7.2.1 Reduced residue systems and Euler's theorem

We start by introducing some useful terminology. Recall that integers a_1, \dots, a_n form a complete residue system modulo n if their remainders when divided by n are a permutation of $0, 1, \dots, n-1$. Considering the totatives² of n instead of $0, 1, \dots, n-1$ naturally yields the following definition.

Definition 7.33. Integers a_1, \dots, a_k form a reduced system of residues mod n (or a reduced residue system mod n) if every integer relatively prime to n is congruent modulo n to exactly one of a_1, \dots, a_k .

Before moving on, let us make the following simple remarks, which are direct consequences of the definition of a reduced system of residues mod n .

Remark 7.34. Clearly a_1, \dots, a_k form a reduced residue system modulo n if and only if their remainders when divided by n are a permutation of the totatives of n . In particular every reduced system of residues mod n has precisely $\varphi(n)$ elements. Moreover, if a_1, a_2, \dots, a_k and b_1, \dots, b_k are reduced systems of residues mod n , then there is a permutation σ of $1, 2, \dots, k$ such that $a_i \equiv b_{\sigma(i)} \pmod{n}$ for all i .

If a_1, \dots, a_n is a complete residue system modulo n , then for any integer a relatively prime to n the numbers aa_1, \dots, aa_n form a complete residue system modulo n . The next proposition establishes a similar result for reduced residue systems.

Proposition 7.35. *If a_1, \dots, a_k is a reduced system of residues mod n and if a is an integer relatively prime to n , then aa_1, aa_2, \dots, aa_k is a reduced system of residues mod n .*

Proof. First, aa_i is relatively prime to n , since a and a_i are so. Next, by remark 7.34, it suffices to prove that aa_1, aa_2, \dots, aa_k are pairwise incongruent mod n . If $aa_i \equiv aa_j \pmod{n}$, by Gauss' lemma we have $a_i \equiv a_j \pmod{n}$, hence $i = j$. The result follows. \square

²Recall that an integer $a \in \{1, 2, \dots, n\}$ is called a totative of n if $\gcd(a, n) = 1$.

We are now ready to state and prove the following important theorem, which generalizes Fermat's little theorem.

Theorem 7.36. (*Euler's theorem*) *If n is a positive integer, then for all integers a relatively prime to n we have*

$$a^{\varphi(n)} \equiv 1 \pmod{n}.$$

Proof. Let a_1, \dots, a_k be a reduced system of residues mod n . By proposition 7.35 the numbers aa_1, \dots, aa_k form a reduced residue system modulo n , thus

$$a_1 a_2 \dots a_k \equiv (aa_1) \cdot (aa_2) \cdot \dots \cdot (aa_k) \pmod{n},$$

by remark 7.34. This congruence can be rewritten as

$$a_1 a_2 \dots a_k (a^{\varphi(n)} - 1) \equiv 0 \pmod{n}.$$

Since $\gcd(n, a_i) = 1$ for all i , it follows that $\gcd(n, a_1 a_2 \dots a_k) = 1$, hence the previous congruence simplifies to $a^{\varphi(n)} - 1 \equiv 0 \pmod{n}$, as needed. \square

We can also prove Euler's theorem as follows. Let p be a prime divisor of n , so that $p-1 \mid \varphi(n)$. By Fermat's little theorem and the lifting the exponent lemma (more precisely theorem 6.22) we have

$$v_p(a^{\varphi(n)} - 1) = v_p\left((a^{p-1})^{\frac{\varphi(n)}{p-1}} - 1\right) \geq v_p(a^{p-1} - 1) + v_p\left(\frac{\varphi(n)}{p-1}\right).$$

We conclude that $v_p(a^{\varphi(n)} - 1) \geq v_p(n)$, since $v_p(a^{p-1} - 1) \geq 1$ and

$$v_p\left(\frac{\varphi(n)}{p-1}\right) = v_p(\varphi(n)) \geq v_p(n) - 1.$$

We illustrate Euler's theorem with some simple examples, the more challenging ones being kept for the next section.

Example 7.37. Prove that for all $a \geq 2$ and $n \geq 1$ we have $n \mid \varphi(a^n - 1)$.

Proof. By Euler's theorem we have $a^{\varphi(a^n - 1)} \equiv 1 \pmod{a^n - 1}$. Thus $a^n - 1 \mid a^{\varphi(a^n - 1)} - 1$. We conclude that $n \mid \varphi(a^n - 1)$ using corollary 3.36. \square

Example 7.38. Prove that $n^2 - 1 \mid 2^{n!} - 1$ for all even integers $n > 0$.

Proof. Since n is even, $n - 1$ and $n + 1$ are relatively prime, thus it suffices to prove that $n - 1$ and $n + 1$ each divide $2^{n!} - 1$. By Euler's theorem $n \pm 1 \mid 2^{\varphi(n \pm 1)} - 1$ and so it is enough to prove that $\varphi(n \pm 1) \mid n!$. This is clear, since $\varphi(n \pm 1) \leq (n \pm 1) - 1 \leq n$. \square

Example 7.39. Let p be prime number. Given an integer a such that $\gcd(a, p!) = 1$, prove that $a^{(p-1)!} - 1$ is divisible by $p!$.

Proof. By Fermat's little theorem $a^{(p-1)!} - 1$ is a multiple of p , thus it suffices to prove that $(p - 1)! \mid a^{(p-1)!} - 1$. If $q < p$ is a prime and $k = v_q((p - 1)!)$, then $\varphi(q^k) = q^{k-1}(q - 1) \mid (p - 1)!$, and Euler's theorem yields the desired result. \square

Example 7.40. Find all positive integers dividing infinitely many numbers in the sequence $1, 11, 111, 1111, \dots$

Proof. Clearly none of the numbers $1, 11, 111, \dots$ is even or a multiple of 5, so any solution of the problem is relatively prime to 2 and 5. Conversely, let n be a positive integer relatively prime to 10. We will prove that for infinitely many k we have $n \mid \frac{10^k - 1}{9}$, or equivalently $9n \mid 10^k - 1$. Simply take $k = M\varphi(9n)$ for any $M \geq 1$ and use Euler's theorem to conclude. \square

We end this section with two more results concerning reduced residue systems. The following theorem relates reduced systems of residues modulo m, n and mn , if m and n are relatively prime positive integers. Note that it immediately implies that Euler's totient function φ is multiplicative, a result that has already been obtained as a consequence of the explicit formula for $\varphi(n)$.

Theorem 7.41. Let a_1, a_2, \dots, a_k be a reduced system of residues mod n and let b_1, b_2, \dots, b_l be a reduced system of residues mod m . If $\gcd(m, n) = 1$, then $(ma_i + nb_j)_{1 \leq i \leq k, 1 \leq j \leq l}$ is a reduced residue system mod mn .

Proof. First, we check that $\gcd(ma_i + nb_j, mn) = 1$ for all i, j . If a prime p divides mn and $ma_i + nb_j$, we may assume that it divides m . Then $p \mid nb_j$ and since $\gcd(m, n) = 1$ we have $p \mid b_j$, contradicting the equality $\gcd(b_j, m) = 1$.

Next, we prove that $ma_i + nb_j$ are pairwise incongruent mod mn . Suppose that $ma_i + nb_j \equiv ma_k + nb_l \pmod{mn}$. Then $nb_j \equiv nb_l \pmod{m}$ and since $\gcd(n, m) = 1$, we must have $b_j \equiv b_l \pmod{m}$, thus $j = l$. We obtain similarly $i = k$.

We prove finally that for any x relatively prime to mn we can find i, j such that $x \equiv ma_i + nb_j \pmod{mn}$. Pick an integer m' such that $mm' \equiv 1 \pmod{n}$ (possible since $\gcd(m, n) = 1$). Then $\gcd(m'x, n) = 1$, hence there is i such that $m'x \equiv a_i \pmod{n}$. Then $x \equiv ma_i \pmod{n}$, and we can write $x = ma_i + nc$ for some integer c . Since $\gcd(x, m) = 1$, we have $\gcd(c, m) = 1$, thus there is j such that $c \equiv b_j \pmod{m}$. Then $x \equiv ma_i + nb_j \pmod{mn}$, as desired. \square

Remark 7.42. The proof of the previous theorem can be shortened using the equality $\varphi(mn) = \varphi(m)\varphi(n)$ combined with remark 7.34. Indeed, using these observations one can simply delete the third paragraph in the above proof. We preferred to give the previous longer proof since it gives an alternative proof of the formula $\varphi(mn) = \varphi(m)\varphi(n)$.

Finally, we describe the remainder modulo n of the product of the elements of a reduced residue system. The next theorem is due to Gauss.

Theorem 7.43. *Let $a_1, a_2, \dots, a_{\varphi(n)}$ be a reduced residue system modulo $n > 2$ and let N be the number of solutions of the congruence $x^2 \equiv 1 \pmod{n}$. Then*

$$\prod_{i=1}^{\varphi(n)} a_i \equiv (-1)^{\frac{N}{2}} \pmod{n}.$$

Proof. If an integer r is relatively prime to n , then so is its inverse r^{-1} modulo n . It follows that we can make pairs of the form (r, r^{-1}) out of the numbers $a_1, \dots, a_{\varphi(n)}$, such that the product of the elements in each pair is 1 modulo n . We have to be a little bit careful, however, since we may have $r = r^{-1}$ for some r , which happens if and only if $r^2 \equiv 1 \pmod{n}$. Hence we can pair all a_i 's but those which satisfy the congruence $x^2 \equiv 1 \pmod{n}$, and so

$$\prod_{i=1}^{\varphi(n)} a_i \equiv \prod_{x^2 \equiv 1 \pmod{n}} x \pmod{n}.$$

It remains to see that the last product is $(-1)^{N/2}$ modulo n . We use a similar argument: if x is a solution of the congruence $x^2 \equiv 1 \pmod{n}$, then so is $-x$, and moreover x is not congruent to $-x$ modulo n (as otherwise n would divide 2, which is excluded by hypothesis). Thus the solutions of the congruence $x^2 \equiv 1 \pmod{n}$ can be partitioned into $N/2$ pairs of the form $(x, -x)$, and the product of the elements in each pair is $-x^2 \equiv -1 \pmod{n}$. Thus

$$\prod_{x^2 \equiv 1 \pmod{n}} x \equiv (-1)^{N/2} \pmod{n}$$

and we are done. \square

Remark 7.44. The precise value of N was found in example 7.11, using the Chinese remainder theorem. We conclude that

$$\prod_{i=1}^k a_i \equiv 1 \pmod{n}$$

unless $n = 4$ or n is of the form p^k or $2p^k$ for some odd prime p and some $k \geq 1$, in which case $\prod_{i=1}^k a_i \equiv -1 \pmod{n}$.

7.2.2 Practicing Euler's theorem

In this section we give several less straightforward examples in which Euler's theorem is the key ingredient. We start with a very short proof of the existence part of the Chinese remainder theorem.

Example 7.45. Prove the existence part of the Chinese remainder theorem using Euler's theorem.

Proof. Let m_1, \dots, m_n be pairwise relatively prime integers and let a_1, \dots, a_n be arbitrary integers. We need to find x such that $x \equiv a_i \pmod{m_i}$ for all i . Simply take

$$x = a_1(m_2 \dots m_n)^{\varphi(m_1)} + a_2(m_1 m_3 \dots m_n)^{\varphi(m_2)} + \dots + a_n(m_1 \dots m_{n-1})^{\varphi(m_n)}.$$

By Euler's theorem x satisfies the desired congruences. \square

We continue with three rather remarkable congruences.

Example 7.46. Prove that for all positive integers n and all integers a

$$\sum_{d|n} \varphi(d) a^{\frac{n}{d}} \equiv 0 \pmod{n}.$$

Proof. Let

$$x_n(a) = \sum_{d|n} \varphi(d) a^{\frac{n}{d}}$$

and let $P(n)$ be the following statement: $n \mid x_n(a)$ for all integers a . First, let us check that if $\gcd(m, n) = 1$ and if $P(m)$ and $P(n)$ are true, then so is $P(mn)$. Let a be an integer. Since $\gcd(m, n) = 1$, it suffices to show that $m \mid x_{mn}(a)$ and $n \mid x_{mn}(a)$. By symmetry, it is enough to prove the divisibility $m \mid x_{mn}(a)$. Note however that since $\gcd(m, n) = 1$ and φ is multiplicative, we have

$$\begin{aligned} x_{mn}(a) &= \sum_{d|mn} \varphi(d) a^{\frac{mn}{d}} = \sum_{e|m, f|n} \varphi(e) \varphi(f) a^{\frac{m}{e} \cdot \frac{n}{f}} \\ &= \sum_{f|n} \varphi(f) \sum_{e|m} \varphi(e) (a^{\frac{n}{f}})^{\frac{m}{e}} = \sum_{f|n} \varphi(f) x_m(a^{\frac{n}{f}}). \end{aligned}$$

Since $P(m)$ holds, each of the numbers $x_m(a^{\frac{n}{f}})$ is a multiple of m , so we are done.

Taking into account the previous discussion, it suffices to prove that $p^n \mid x_{p^n}(a)$ for all a , $n \geq 1$ and primes p . Note however that

$$\begin{aligned} x_{p^n}(a) &= a^{p^n} + (p-1)a^{p^{n-1}} + p(p-1)a^{p^{n-2}} + \dots + p^{n-1}(p-1)a \\ &= a^{p^n} - a^{p^{n-1}} + p(a^{p^{n-1}} + (p-1)a^{p^{n-2}} + \dots + p^{n-2}(p-1)a) \\ &= a^{p^n} - a^{p^{n-1}} + px_{p^{n-1}}(a). \end{aligned}$$

Thus, arguing by induction on n , it suffices to prove that $p \mid x_p(a)$ (which is equivalent to $a^p \equiv a \pmod{p}$, i.e. Fermat's little theorem) and $p^n \mid a^{p^n} - a^{p^{n-1}}$. This last divisibility is clear if $p \mid a$, and otherwise it follows from Euler's theorem. \square

Example 7.47. Prove that, for all positive integers n and all integers a ,

$$n \mid \sum_{i=1}^n a^{\gcd(i,n)}$$

Proof. If d is a positive divisor of n , then the integers $i \in \{1, 2, \dots, n\}$ for which $\gcd(i, n) = d$ are precisely the numbers dj with j a totative of $\frac{n}{d}$, thus there are $\varphi(\frac{n}{d})$ such integers i . We deduce that

$$\sum_{i=1}^n a^{\gcd(i,n)} = \sum_{d|n} \varphi\left(\frac{n}{d}\right) a^d = \sum_{d|n} \varphi(d) a^{\frac{n}{d}}$$

and the result follows then from example 7.46. \square

Example 7.48. (IMO Shortlist 1987) Let $(a_n)_{n \geq 1}$ be a sequence of integers satisfying

$$\sum_{d|n} a_d = 2^n$$

for all n . Prove that n divides a_n for all n .

Proof. It is immediate to check the property for $n = 1$ and $n = 2$. Assume, by strong induction, that $n > 2$ and that a_k is divisible by k for all $k < n$. It suffices to prove that if p is a prime and $m = v_p(n)$, then p^m divides a_n . By hypothesis

$$a_n = 2^n - \sum_{d|n, d < n} a_d.$$

If $d < n$ is a divisor of n for which $p^m \mid d$, then $p^m \mid d \mid a_d$. Thus

$$a_n \equiv 2^n - \sum_{d|n/p} a_d = 2^n - 2^{n/p} \pmod{p^m}.$$

It suffices to prove that $2^n - 2^{n/p}$ is a multiple of p^m . If $p = 2$, this is clear, since $n/p \geq m$ (because p^m divides n , we have $n/p \geq p^{m-1} = 2^{m-1} \geq m$). So assume that $p > 2$. By Euler's theorem, it is enough to check that $n - \frac{n}{p}$ is a multiple of $\varphi(p^m) = p^{m-1}(p-1)$, or equivalently that n is a multiple of p^m , which holds by definition of m . \square

The next examples have a more combinatorial and constructive nature.

Example 7.49. Let a_1, \dots, a_n be rational numbers such that $a_1^k + a_2^k + \dots + a_n^k$ is an integer for all $k \geq 1$. Prove that a_1, \dots, a_n are integers.

Proof. Let d be the product of the denominators of a_1, \dots, a_n and write $x_i = da_i$, then x_1, \dots, x_n are integers and by assumption $d^k \mid x_1^k + \dots + x_n^k$ for all $k \geq 1$. We want to prove that $d \mid x_i$ for all i . Using the prime factorization of d , we may assume that d is a power of a prime p , say $d = p^j$. By an immediate induction on j , we may assume that $j = 1$. Thus $p^k \mid x_1^k + \dots + x_n^k$ for all $k \geq 1$ and we want to prove that $p \mid x_1, \dots, x_n$. Assume that this is not the case and let I be the set of those $i \in \{1, \dots, n\}$ for which p does not divide x_i . Using Euler's theorem we obtain

$$x_1^{\varphi(p^k)} + \dots + x_n^{\varphi(p^k)} \equiv |I| \pmod{p^k}.$$

On the other hand, by assumption $p^{\varphi(p^k)}$ (and thus p^k) divides the left-hand side. We deduce that $p^k \mid |I|$ and since $k \geq 1$ was arbitrary, it follows that $|I| = 0$, a contradiction. The result follows. \square

Remark 7.50. The conclusion is trivially false without the assumption that a_1, \dots, a_n are rational numbers (consider for instance $a_1 = 1 + \sqrt{2}$ and $a_2 = 1 - \sqrt{2}$). The most general result (whose proof is outside the scope of this book) is the following: for complex numbers a_1, \dots, a_n the numbers $a_1^k + \dots + a_n^k$ are integers for all $k \geq 1$ if and only if $\prod_{i=1}^n (X - a_i)$ has integer coefficients.

Example 7.51. (China TST 2006) Prove that for any positive integers m, n there is a positive integer k such that $2^k - m$ has at least n different prime divisors.

Proof. By replacing m with its largest odd divisor, we may assume that m is odd. Let $\omega(x)$ be the number of different prime divisors of $x > 1$. It suffices to prove that if $2^k - m > 1$ then we can find $l > k$ such that $\omega(2^l - m) > \omega(2^k - m)$. Let $2^k - m = p_1^{\alpha_1} \dots p_N^{\alpha_N}$ be the prime factorization of $2^k - m$ and note that $p_i > 2$ for all i , since m is odd. Choose $l = k + \prod_{i=1}^N \varphi(p_i^{\alpha_i+1})$ and note that by Euler's theorem we have

$$2^l - m \equiv 2^k - m \pmod{p_i^{\alpha_i+1}}$$

in particular $v_{p_i}(2^l - m) = \alpha_i = v_{p_i}(2^k - m)$ for all $1 \leq i \leq N$. Since $2^l - m > 2^k - m$, it follows that $2^l - m$ must have a prime factor different from p_1, \dots, p_N , thus $\omega(2^l - m) > \omega(2^k - m)$ and we are done. \square

Example 7.52. Let y be a positive integer. Prove that there are infinitely many primes p such that $p \equiv -1 \pmod{4}$ and $p \mid 2^n y + 1$ for some positive integer n .

Proof. We may assume that y is odd, so that $2y + 1 \equiv -1 \pmod{4}$. Suppose that p_1, \dots, p_k are all primes of the form $4m + 3$ which divide at least one of the numbers $2y + 1, 4y + 1, 8y + 1, \dots$. Set $n = \varphi((2y + 1)p_1 \dots p_k) + 1$. By Euler's theorem we have

$$2^n y + 1 \equiv 2y + 1 \pmod{(2y + 1)p_1 \dots p_k}.$$

Hence we can write $2^n y + 1 = (2y + 1)(sp_1 \dots p_k + 1)$ for some positive integer s . Since $2^n y + 1 \equiv 1 \pmod{4}$ and $2y + 1 \equiv 3 \pmod{4}$, we must have $sp_1 \dots p_k + 1 \equiv -1 \pmod{4}$, hence there is a prime $q \equiv -1 \pmod{4}$ such that $q \mid sp_1 \dots p_k + 1$. But then $q \mid 2^n y + 1$, so $q \in \{p_1, \dots, p_k\}$, obviously impossible. The result follows. \square

Example 7.53. (IMO Shortlist 2012) Let x and y be positive integers. If $x^{2^n} - 1$ is divisible by $2^n y + 1$ for every positive integer n , prove that $x = 1$.

Proof. Suppose that there is a prime q such that $q \mid 2^n y + 1$ and $q \equiv -1 \pmod{4}$, then we get that $q \mid x^{2^n} - 1 = (x - 1)(x + 1)(x^2 + 1)(x^4 + 1) \dots (x^{2^{n-1}} + 1)$. But q cannot divide $x^{2^m} + 1$ for any positive integer m (see corollary 5.28), so $q \mid x^2 - 1$. We conclude using the previous example. \square

Example 7.54. Let a_1, \dots, a_n be positive integers, not all equal. Prove that the set of prime numbers dividing at least one of the numbers $a_1^k + a_2^k + \dots + a_n^k$ with $k \geq 1$ is infinite.

Proof. We may assume that $\gcd(a_1, \dots, a_n) = 1$. Write $f(k) = a_1^k + \dots + a_n^k$ for $k \geq 1$ and suppose that all prime divisors of $f(1), f(2), \dots$ belong to $\{p_1, \dots, p_N\}$ for some primes p_1, \dots, p_N and some $N \geq 1$. For each $1 \leq i \leq N$, let b_i be the number of terms of the sequence a_1, \dots, a_n which are not divisible by p_i . Since $\gcd(a_1, \dots, a_n) = 1$, we have $b_i \geq 1$ for all $1 \leq i \leq N$.

Note that for

$$k = 2 \prod_{i=1}^N \varphi(p_i^{1+v_{p_i}(b_i)})$$

we have $f(k) \equiv b_i \pmod{p_i^{v_p(b_i)+1}}$ for all $1 \leq i \leq N$, since for any $1 \leq j \leq n$ we have $a_j^k \equiv 1 \pmod{p_i^{v_p(b_i)+1}}$ if p_i does not divide a_j (by Euler's theorem) and $a_j^k \equiv 0 \pmod{p_i^{v_p(b_i)+1}}$ otherwise (since $k > 1 + v_p(b_i)$). Therefore $v_{p_i}(f(k)) = v_{p_i}(b_i)$ for all i and since all prime divisors of $f(k)$ belong to $\{p_1, p_2, \dots, p_N\}$, we conclude that $f(k) = p_1^{v_{p_1}(b_1)} p_2^{v_{p_2}(b_2)} \dots p_N^{v_{p_N}(b_N)}$. Since $\max(a_1, \dots, a_n) \geq 2$, we have

$$f(k) \geq 2^k > k > \prod_{i=1}^N p_i^{v_p(b_i)},$$

a contradiction. The result follows. \square

Example 7.55. (USA TST 2007) Are there integers $a, b \geq 1$ such that a does not divide $b^n - n$ for all $n \geq 1$?

Proof. The answer is negative. We will prove by strong induction on a the following: for all $b \geq 1$ there are infinitely many n such that $a \mid b^n - n$. This is clear for $a = 1$, so assume that it holds up to $a - 1$ and let us prove it for a . Since $\varphi(a) < a$, the inductive hypothesis yields the existence of infinitely many n such that $\varphi(a) \mid b^n - n$. We claim that if $\varphi(a) \mid b^n - n$ and n is big enough, then $a \mid b^{b^n} - b^n$, which is enough to conclude. To prove the claim, write $b^n - n = c\varphi(a)$, then

$$b^{b^n} - b^n = b^{n+c\varphi(a)} - b^n = b^n((b^c)^{\varphi(a)} - 1).$$

Take now any prime factor p of a and let $k = v_p(a)$. If p does not divide b , then Euler's theorem gives $p^k \mid (b^c)^{\varphi(p^k)} - 1 \mid (b^c)^{\varphi(a)} - 1$. On the other hand, if $p \mid b$ and $n \geq k$, then certainly $p^k \mid b^n$. Thus if $n \geq \max_{p \mid a} v_p(a)$, then $a \mid b^n((b^c)^{\varphi(a)} - 1)$, finishing the proof. \square

Example 7.56. (Russia 2004) Is there an integer $n > 10^{1000}$ which is not divisible by 10 such that one can exchange two distinct non-zero digits in its decimal representation without changing the set of prime divisors of n ?

Proof. Yes, there is such a number, actually there are infinitely many of them! For each positive integer k let

$$n_k = 13 \cdot \frac{10^{360k} - 1}{9} = 144 \dots 43.$$

Exchanging the digits 1 and 3 we obtain the number $344 \dots 41 = 31 \cdot \frac{10^{360k} - 1}{9}$, which has the same prime divisors since $10^{360k} - 1$ is divisible by both 13 and 31 by Euler's theorem (because $360 = \varphi(13 \cdot 31)$). \square

7.3 Order modulo n

7.3.1 Elementary properties and examples

Let n be a positive integer and let a be an integer relatively prime to n . By Euler's theorem there are infinitely many positive integers k such that $a^k \equiv 1 \pmod{n}$, for instance all multiples of $\varphi(n)$. In this section we study in more detail the congruence $a^x \equiv 1 \pmod{n}$. We will see that all solutions of this congruence are determined by the smallest positive solution. The following definition is therefore rather natural.

Definition 7.57. If n is a positive integer and a is an integer relatively prime to n , the smallest positive solution of the congruence $a^x \equiv 1 \pmod{n}$ is called the order of a modulo n and denoted $\text{ord}_n(a)$.

Note that $\text{ord}_n(a)$ is not defined when a is not relatively prime to n . Also, the sequence of remainders mod n of the numbers $1, a, a^2, \dots$ is periodic with (minimal) period $\text{ord}_n(a)$. This follows from the fact that $a^i \equiv a^{i+j} \pmod{n}$ is equivalent (by Gauss' lemma) to $a^j \equiv 1 \pmod{n}$ for all positive integers i, j . For instance, consider $a = 3$ and $n = 17$, then the sequence of remainders of $1, a, a^2, \dots$ when divided by n is

$$1, 3, 9, 10, 13, 5, 15, 11, 16, 14, 8, 7, 4, 12, 2, 6, 1, 3, 9, \dots$$

and the length of the period is 16 hence $\text{ord}_{17}(3) = 16$.

The following fundamental theorem summarizes the most important properties of $\text{ord}_n(a)$.

Theorem 7.58. *Let a be an integer relatively prime to $n > 1$.*

a) The positive solutions of the congruence $a^x \equiv 1 \pmod{n}$ are exactly the multiples of $\text{ord}_n(a)$.

b) $\text{ord}_n(a)$ divides $\varphi(n)$.

Proof. Note that b) follows from a) and Euler's theorem, so it suffices to prove part a). Let $d = \text{ord}_n(a)$. Since $a^d \equiv 1 \pmod{n}$ we have $a^{md} \equiv 1 \pmod{n}$ for all $m \geq 1$, so all multiples of d are solutions of the congruence $a^x \equiv 1 \pmod{n}$. Conversely, let $k > 0$ be such that $a^k \equiv 1 \pmod{n}$ and consider the Euclidean division $k = q \cdot d + r$, with $0 \leq r < d$. Then

$$1 \equiv a^k \equiv a^{qd} \cdot a^r \equiv a^r \pmod{n},$$

thus $a^r \equiv 1 \pmod{n}$. Since $r < d$, the minimality of d forces $r = 0$ and so $d \mid k$, finishing the proof. \square

Part b) of the previous theorem is very useful especially when $\varphi(n)$ has a simple form. Here are a few relatively simple examples that illustrate this result (the reader will find more challenging examples in the next section).

Example 7.59. Determine $\text{ord}_n(a)$ in the following cases:

a) $a = 2$ and $n \in \{7, 11, 15\}$.

b) $a = 5$ and $n \in \{7, 11, 23\}$.

Proof. In all cases we let $d = \text{ord}_n(a)$ and we use that $d \mid \varphi(n)$.

a) Suppose that $n = 7$, so $\varphi(7) = 6$ and $d \mid 6$. Checking successively divisors of 6 yields $d = 3$. Suppose that $n = 11$, then $d \mid 10$. Again, checking the divisors 1, 2, 5, 10 of 10 yields $d = 10$. For $n = 15$ we have $\varphi(n) = 8$ and $d \mid 8$. Since $2^4 \equiv 1 \pmod{15}$ and 2^2 is not congruent to 1 mod 15 we deduce that $d = 4$ in this case.

b) For $n = 7$ we have $d \mid 6$ and since 7 does not divide $5^2 - 1$ and $5^3 - 1$ we deduce that $d = 6$. For $n = 11$ we have $d \mid 10$ and 11 does not divide $5^2 - 1$. Next $5^5 \equiv 25 \cdot 125 \equiv 3 \cdot 4 \equiv 1 \pmod{11}$, so $d = 5$. Finally, for $n = 23$ we have $d \mid 22$ and 23 does not divide $5^2 - 1$. Also,

$$5^{11} \equiv 5 \cdot 25^5 \equiv 5 \cdot 2^5 \equiv 5 \cdot 9 \equiv -1 \pmod{23}$$

hence $d = 22$. \square

Example 7.60. Let n be an integer greater than 1.

a) Compute $\text{ord}_{2^n}(5)$ and prove that

$$1, 5, 5^2, \dots, 5^{2^{n-2}-1}, -1, -5, \dots, -5^{2^{n-2}-1}$$

form a reduced residue system modulo 2^n .

b) Prove that for any $a \equiv 1 \pmod{4}$ there is a unique $i \in \{0, 1, \dots, 2^{n-2}-1\}$ such that $a \equiv 5^i \pmod{2^n}$, and for any $a \equiv -1 \pmod{4}$ there is a unique $i \in \{0, 1, \dots, 2^{n-2}-1\}$ such that $a \equiv -5^i \pmod{2^n}$.

Proof. a) Let $d = \text{ord}_{2^n}(5)$, then $d \mid \varphi(2^n) = 2^{n-1}$, so $d = 2^k$ for some $0 \leq k < n$. Thus we need to find the smallest $k \geq 0$ for which $2^n \mid 5^{2^k} - 1$, i.e. such that $v_2(5^{2^k} - 1) \geq n$. Using either (and preferably!) the factorization

$$5^{2^k} - 1 = (5 - 1)(5 + 1)(5^2 + 1) \dots (5^{2^{k-1}} + 1)$$

or the lifting the exponent lemma, we obtain $v_2(5^{2^k} - 1) = k + 2$. Thus the inequality $v_2(5^{2^k} - 1) \geq n$ is equivalent to $k \geq n - 2$ and so $d = 2^{n-2}$.

Since $\varphi(2^n) = 2^{n-1}$, any reduced residue system modulo 2^n has 2^{n-1} elements. It suffices therefore to prove that the numbers $1, 5, 5^2, \dots, 5^{2^{n-2}-1}, -1, -5, \dots, -5^{2^{n-2}-1}$ give different remainders when divided by 2^n . Since $\text{ord}_{2^n}(5) = 2^{n-2}$, the numbers $1, 5, 5^2, \dots, 5^{2^{n-2}-1}$ give different remainders mod 2^n , and similarly for the numbers $-1, -5, \dots, -5^{2^{n-2}-1}$. Finally, we cannot have $5^i \equiv -5^j \pmod{2^n}$, for some $0 \leq i, j \leq 2^{n-2} - 1$, since this would imply that $1 \equiv -1 \pmod{4}$, a contradiction. The result follows.

b) This is an immediate consequence of part a) and of the fact that $5^k \equiv 1 \pmod{4}$ for all k , while $-5^k \equiv 3 \pmod{4}$ for all k . \square

The result established in part a) of the next example is very important.

Example 7.61. (Lucas, 1878) Let $n > 1$ be an integer and let p be a prime divisor of $F_n = 2^{2^n} + 1$.

a) Prove that the order of 2 modulo p is 2^{n+1} and deduce that $2^{n+1} \mid p - 1$.

b) Prove that $a = 2^{2^{n-2}}(2^{2^{n-1}} - 1)$ has order 2^{n+2} modulo p and deduce that $2^{n+2} \mid p - 1$.

c) Prove that if $p^2 \mid F_n$, then $p^2 \mid 2^{p-1} - 1$.

d) Prove that $p \mid 2^{\frac{p-1}{2}} - 1$ and deduce a new proof of the fact that $2^{n+2} \mid p - 1$.

Proof. a) Let d be the order of 2 modulo p . Since $2^{2^n} \equiv -1 \pmod{p}$, we have $2^{2^{n+1}} \equiv 1 \pmod{p}$, thus d divides 2^{n+1} . If d divided 2^n , then $2^{2^n} \equiv 1 \pmod{p}$ and since $2^{2^n} \equiv -1 \pmod{p}$, we would obtain $2 \equiv 0 \pmod{p}$, a plain contradiction. Thus d divides 2^{n+1} and does not divide 2^n , which means that $d = 2^{n+1}$. Since d divides $\varphi(p) = p - 1$, we are done.

b) Note that

$$a^2 = 2^{2^{n-1}}(2^{2^n} - 2 \cdot 2^{2^{n-1}} + 1) \equiv -2 \cdot 2^{2^{n-1}+2^{n-1}} \equiv -2(-1) = 2 \pmod{p},$$

since $p \mid 2^{2^n} + 1$. We deduce that $a^{2^{n+1}} \equiv 2^{2^n} \equiv -1 \pmod{p}$. Arguing as in a) we deduce that the order of a modulo p divides 2^{n+2} and does not divide 2^{n+1} , thus it equals 2^{n+2} . Since the order divides $\varphi(p) = p - 1$, we deduce that $2^{n+2} \mid p - 1$.

c) Since $p \equiv 1 \pmod{2^{n+1}}$ by part a), we obtain

$$p^2 \mid F_n \mid 2^{2^{n+1}} - 1 \mid 2^{p-1} - 1,$$

as needed.

d) The divisibility $p \mid 2^{\frac{p-1}{2}} - 1$ is equivalent, by Euler's criterion (theorem 5.99) to $\left(\frac{2}{p}\right) = 1$, which is equivalent (by theorem 5.125) to $p \equiv \pm 1 \pmod{8}$. Since $p \equiv 1 \pmod{8}$ by part a), we obtain $p \mid 2^{\frac{p-1}{2}} - 1$. Next, since the order of 2 modulo p is 2^{n+1} (again by part a)) and since $2^{\frac{p-1}{2}} \equiv 1 \pmod{p}$, we obtain $2^{n+1} \mid \frac{p-1}{2}$ and so $p \equiv 1 \pmod{2^{n+2}}$. \square

Remark 7.62. The only known primes p satisfying $p^2 \mid 2^{p-1} - 1$ are 1093 and 3511, discovered in 1913 and 1922 by Meissner and Beeger. These primes are called Wieferich primes and it is an open problem whether there are infinitely many such primes. Note that 1093 and 3511 cannot divide any Fermat number, since 2^7 does not divide 1092 or 3510, while by Lucas' theorem any prime factor of $2^{2^n} + 1$ with $n \geq 5$ is congruent to 1 modulo 2^7 . Therefore not a single Fermat number which is not squarefree is currently known!

Combining the next example and the previous remark shows that $2^p - 1$ is quite likely squarefree when p is a prime (again, no counterexample to this assertion is known).

Example 7.63. Suppose that p, q are primes and $p^2 \mid 2^q - 1$. Prove that $2^{p-1} \equiv 1 \pmod{p^2}$.

Proof. Let d be the order of 2 modulo p^2 . Then $d \mid \varphi(p^2) = p(p-1)$ and the hypothesis yields $d \mid q$. Clearly $d \neq 1$, thus necessarily $d = q$ and so $q \mid p(p-1)$. If $q = p$, we obtain $p \mid 2^p - 1$, clearly impossible by Fermat's little theorem. Thus $q \mid p-1$. But then $p^2 \mid 2^q - 1 \mid 2^{p-1} - 1$, as needed. \square

Example 7.64. Let $n > 1$ be an integer such that $a = 2^n + 1$ is pseudo-prime, i.e. $a \mid 2^a - 2$. Prove that n is a power of 2.

Proof. The hypothesis yields $2^n + 1 \mid 2^{2^n} - 1$. Let d be the order of 2 modulo $2^n + 1$. Since $2^{2^n} \equiv 1 \pmod{2^n + 1}$, we have $d \mid 2^n$, so d is a power of 2. On the other hand, $2^n \equiv -1 \pmod{2^n + 1}$, thus $2^{2^n} \equiv 1 \pmod{2^n + 1}$ and $d \mid 2n$. If $d \neq 2n$, then $d \leq n$ and so $2^n + 1 \leq 2^d - 1 < 2^n$, impossible. Thus $d = 2n$ and since d is a power of 2, it follows that n is a power of 2. \square

Example 7.65. (Kvant M 1355) Let n be a positive integer such that $2^{2n} + 2^n + 1$ is a prime. Prove that this prime is a divisor of $2^{2^{2n}+1} - 1$.

Proof. Let $p = 2^{2n} + 2^n + 1$ and note that $p \mid 2^{3n} - 1$. Thus in order to show that $p \mid 2^{2^{2n}+1} - 1$ it suffices to prove that $3n \mid 2^n + 1$. Let $d = \text{ord}_p(2)$. Since $2^{3n} \equiv 1 \pmod{p}$ we have $d \mid 3n$. Next, we have $d > 2n > \frac{3n}{2}$ since $2^d \equiv 1 \pmod{p}$ (thus $2^d > p > 2^{2n}$), which combined with $d \mid 3n$ yields $d = 3n$. Since $d \mid p - 1$, we conclude that $3n \mid p - 1 = 2^n(2^n + 1)$. Finally, note that n is odd (if n is even then $p > 3$ and $p \equiv 0 \pmod{3}$, a contradiction) hence $\gcd(3n, 2^n) = 1$ and so $3n \mid 2^n + 1$, as desired. \square

We present a few more theoretical results that can be very helpful when dealing with orders modulo n . The first one says that if one knows how to compute $\text{ord}_n(a)$, then one can also easily compute $\text{ord}_n(a^k)$ for all $k \geq 1$.

Proposition 7.66. *Let a, n be relatively prime integers, with $n > 1$, and let $d = \text{ord}_n(a)$. Then for any positive integer k*

$$\text{ord}_n(a^k) = \frac{d}{\gcd(d, k)}.$$

In particular

a) *We have $\text{ord}_n(a^k) = d$ if and only if $\gcd(d, k) = 1$.*

b) *If $k \mid d$, then $\text{ord}_n(a^k) = \frac{d}{k}$.*

Proof. Let $m = \gcd(d, k)$ and write $d = md_1, k = mk_1$ with $\gcd(d_1, k_1) = 1$. Setting $t = \text{ord}_n(a^k)$, we have

$$(a^k)^{d_1} = a^{mk_1d_1} = (a^d)^{k_1} \equiv 1 \pmod{n},$$

hence $t \mid d_1$. On the other hand, since $a^{kt} = (a^k)^t \equiv 1 \pmod{n}$ we must have $d \mid kt$, thus $d_1 \mid k_1t$. As d_1 and k_1 are relatively prime, we have $d_1 \mid t$. We conclude that $t = d_1$, as desired. \square

The next result reduces the computation of $\text{ord}_n(a)$ to the case when n is a power of a prime.

Proposition 7.67. *Let a, n be relatively prime integers, with $n > 1$. Let $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$ be the prime factorization of n . Then*

$$\text{ord}_n(a) = \text{lcm}(\text{ord}_{p_1^{\alpha_1}}(a), \dots, \text{ord}_{p_k^{\alpha_k}}(a)).$$

Proof. To simplify notations, let $d = \text{ord}_n(a)$ and $d_i = \text{ord}_{p_i^{\alpha_i}}(a)$ for $1 \leq i \leq k$. Finally, let $M = \text{lcm}(d_1, \dots, d_k)$. Since $a^{d_i} \equiv 1 \pmod{p_i^{\alpha_i}}$ and $d_i \mid M$, we have $a^M \equiv 1 \pmod{p_i^{\alpha_i}}$ for all $1 \leq i \leq k$ and so $a^M \equiv 1 \pmod{n}$. It follows that $d \mid M$. On the other hand $a^d \equiv 1 \pmod{n}$, thus $a^d \equiv 1 \pmod{p_i^{\alpha_i}}$ for all $1 \leq i \leq k$ and so $d_i \mid d$ for all $1 \leq i \leq k$. It follows that $M \mid d$ and then $d = M$, as desired. \square

Finally, the following rather technical result reduces the computation of $\text{ord}_{p^k}(a)$ to computing $\text{ord}_p(a)$ and $v_p(a^{\text{ord}_p(a)} - 1)$. It is a simple consequence of the lifting the exponent lemma (which has already been used when discussing

example 7.60). We strongly advise the reader to repeat the proof every time he needs to compute expressions of the form $\text{ord}_{p^k}(a)$, instead of memorizing the rather messy formulae.

Proposition 7.68. *Let p be a prime, α a positive integer and $a > 1$ an integer relatively prime to p . Let $d = \text{ord}_p(a)$ and let $v = v_p(a^d - 1) \geq 1$.*

a) Suppose that $p > 2$. If $v \geq \alpha$ then $\text{ord}_{p^\alpha}(a) = d$, otherwise

$$\text{ord}_{p^\alpha}(a) = d \cdot p^{\alpha-v}.$$

In particular, if $v_p(a^{\text{ord}_p(a)} - 1) = 1$, then

$$\text{ord}_{p^\alpha}(a) = \text{ord}_p(a) \cdot p^{\alpha-1}.$$

b) Suppose that $p = 2$ and $\alpha > 1$. If $a \equiv 1 \pmod{2^\alpha}$ then $\text{ord}_{2^\alpha}(a) = 1$ and if $a \equiv -1 \pmod{2^\alpha}$ then $\text{ord}_{2^\alpha}(a) = 2$. In all other cases

$$\text{ord}_{2^\alpha}(a) = 2^{\alpha-v_2\left(\frac{a^2-1}{2}\right)}.$$

Proof. a) Let $k = \text{ord}_{p^\alpha}(a)$. Then $p^\alpha \mid a^k - 1$, thus $p \mid a^k - 1$ and so $d \mid k$. Clearly, if $v \geq \alpha$ then $p^\alpha \mid a^d - 1$ thus $k \mid d$ and then $k = d$. Assume now that $v < \alpha$ and write $k = dl$ for some positive integer l . Since $p^\alpha \mid a^k - 1$ and $p \mid a^d - 1$, the lifting the exponent lemma yields

$$\alpha \leq v_p(a^k - 1) = v_p(a^{dl} - 1) = v_p(a^d - 1) + v_p(l) = v + v_p(l).$$

It follows that $v_p(l) \geq \alpha - v$ and so $p^{\alpha-v} \mid l$, thus $d \cdot p^{\alpha-v} \mid k$. Conversely, the same calculation shows that $p^\alpha \mid a^{d \cdot p^{\alpha-v}} - 1$ and so $k \mid d \cdot p^{\alpha-v}$. The result follows.

b) The first part is clear, so assume that a is not congruent to ± 1 modulo 2^α , so that $\alpha > v_2\left(\frac{a^2-1}{2}\right)$. Let $k = \text{ord}_{2^\alpha}(a)$, then $k \mid 2^{\alpha-1}$ and so $k = 2^r$ for some $r \geq 0$. Moreover, using the lifting the exponent lemma yields

$$\alpha \leq v_2(a^k - 1) = v_2\left(\frac{a^2 - 1}{2}\right) + r,$$

thus $r \geq \alpha - v_2\left(\frac{a^2-1}{2}\right)$ and $2^{\alpha-v_2\left(\frac{a^2-1}{2}\right)} \mid k$. A similar computation shows that for $n = 2^{\alpha-v_2\left(\frac{a^2-1}{2}\right)}$ we have $a^n \equiv 1 \pmod{2^\alpha}$ and the result follows. \square

Remark 7.69. If $v_p(a^{\text{ord}_p(a)} - 1) > 1$, then $p^2 \mid a^{p-1} - 1$ (since $\text{ord}_p(a) \mid p-1$, hence $a^{\text{ord}_p(a)} - 1 \mid a^{p-1} - 1$). For any a , this happens for very few primes p (see remark 7.62 for the case $a = 2$).

Example 7.70. Prove that if n is a positive integer, then the order of 2 modulo 5^n is equal to $4 \cdot 5^{n-1}$.

Proof. We clearly have $\text{ord}_5(2) = 4$ and $v_5(2^4 - 1) = 1$. Using part a) of the proposition, we obtain

$$\text{ord}_{5^n}(2) = 4 \cdot 5^{n-1},$$

as needed. \square

Example 7.71. Prove that if p is an odd prime and n is a positive integer, then the order of $1+p$ modulo p^n is p^{n-1} .

Proof. The order of $1+p$ modulo p is clearly 1 and $v_p((1+p)^1 - 1) = 1$. Thus the result follows directly from proposition 7.68. \square

Example 7.72. (China Western Olympiad 2010) Let m, k be nonnegative integers and suppose that $p = 2^{2^m} + 1$ is a prime number. Prove that the order of 2 modulo p^{k+1} is $2^{m+1}p^k$.

Proof. For $k = 0$ we need to prove that $\text{ord}_p(2) = 2^{m+1}$, which has already been established (see example 7.61). Next, using part a) of proposition 7.68 we obtain

$$\text{ord}_{p^{k+1}}(2) = 2^{m+1} \cdot 2^{k+1-v_p(2^{2^{m+1}}-1)}.$$

Since we clearly have $v_p(2^{2^{m+1}} - 1) = v_p((p-2)p) = 1$, we obtain

$$\text{ord}_{p^{k+1}}(2) = 2^{m+1}p^k,$$

as desired. \square

We end this rather long section with a rather surprising and very useful connection between order and decimal expansions. This requires some preliminary discussion. If $x \in [0, 1)$ is a real number, then we can attach to x a sequence of digits $a_1, a_2, \dots \in \{0, 1, \dots, 9\}$ as follows: define $a_1 = [10x]$ and $b_1 = 10x - a_1 \in [0, 1)$, then $a_2 = [10b_1]$ and $b_2 = 10b_1 - a_2$, and so on. It is an easy exercise to check that for all $n \geq 1$ we have

$$0 \leq x - \left(\frac{a_1}{10} + \frac{a_2}{10^2} + \dots + \frac{a_n}{10^n} \right) < \frac{1}{10^n},$$

thus the sequence of rational numbers $\left(\frac{a_1}{10} + \frac{a_2}{10^2} + \dots + \frac{a_n}{10^n} \right)_{n \geq 1}$ approximates x to arbitrary precision. The expression $0.a_1a_2\dots$ is called the decimal expansion of x . If x is an arbitrary real number, we can write $|x| = \pm x = N + z$ with N a non-negative integer and $z \in [0, 1)$. If $N = b_k \cdot 10^k + \dots + b_1 \cdot 10 + b_0$ is the base ten expansion of N and $0.a_1a_2\dots$ is the decimal expansion of z , we call $\pm b_k\dots b_0.a_1a_2\dots$ the decimal expansion of x . We say that this decimal expansion is periodic if the sequence $(a_n)_{n \geq 1}$ is eventually periodic, i.e. there is $T \geq 1$ such that for all sufficiently large n we have $a_n = a_{n+T}$. The decimal expansion is called purely periodic if it is periodic starting from the decimal point, i.e. there is $T \geq 1$ such that $a_n = a_{n+T}$ for all $n \geq 1$.

Theorem 7.73. *Let x be a real number.*

- a) The decimal expansion of x is periodic if and only if x is rational.*
- b) The decimal expansion of x is purely periodic if and only if x is rational and the denominator of x (when written in lowest terms) is relatively prime to 10.*
- c) If x is rational and the denominator of x is of the form $2^u 5^v q$ with $\gcd(q, 10) = 1$, then the minimal length of a period of the decimal expansion of x is the order of $10 \bmod q$.*

Proof. Suppose that the decimal expansion of x is periodic, say

$$x = n.a_1\dots a_s b_1\dots b_k b_1\dots b_k b_1\dots b_k \dots$$

for some integer n and some digits $a_1, \dots, a_s, b_1, \dots, b_k$. Then

$$x = n + \frac{\overline{a_1\dots a_s}}{10^s} + \frac{\overline{b_1\dots b_k}}{10^{k+s}} + \frac{\overline{b_1\dots b_k}}{10^{2k+s}} + \dots$$

thus

$$x = n + \frac{\overline{a_1 \dots a_s}}{10^s} + \frac{\overline{b_1 \dots b_k}}{10^s(10^k - 1)},$$

which is clearly a rational number. Moreover, this formula shows that if the decimal expansion of x is purely periodic (thus we can take $s = 0$), then x is a rational number whose denominator is relatively prime to 10 (since the denominator divides $10^k - 1$). This already shows one implication in both a) and b).

Let now x be a rational number and choose a large enough s so that the denominator of $10^s x$ is relatively prime to 10. Using the Euclidean division we can write

$$10^s x = y + \frac{z}{q}$$

for some integers y, z, q with $0 \leq z < q$. Let $k = \text{ord}_q(10)$ be the order of 10 modulo q , thus

$$10^s x = y + \frac{z \cdot \frac{10^k - 1}{q}}{10^k - 1} = y + \frac{N}{10^k - 1}$$

where $0 \leq N < 10^k - 1$ is some integer. Writing

$$y = 10^s n + \overline{a_1 \dots a_s}, \quad N = \overline{b_1 \dots b_k}$$

for some integer n and some digits a_i, b_j , we obtain

$$x = n + \frac{\overline{a_1 \dots a_s}}{10^s} + \frac{\overline{b_1 \dots b_k}}{10^s(10^k - 1)} = n.a_1 \dots a_s b_1 \dots b_k b_1 \dots b_k b_1 \dots b_k \dots$$

This shows that the decimal expansion of x is periodic, a period being given by $k = \text{ord}_q(10)$. Moreover, if the denominator of x is relatively prime to 10, then we can take $s = 0$ in the previous argument and deduce that the decimal expansion of x is purely periodic. This finishes the proof of parts a) and b) of the theorem, and it also shows that the minimal length of a period of the decimal expansion cannot exceed $k = \text{ord}_q(10)$. On the other hand, if k is some period of the decimal expansion of x , then the previous arguments show that we can write

$$10^s x = A + \frac{B}{10^k - 1}$$

for some integers s, A, B . If the denominator of x is $2^u 5^v q$ with $\gcd(q, 10) = 1$, this shows that $q \mid 10^k - 1$ and so $\text{ord}_q(10) \mid k$. Thus the period k must be at least $\text{ord}_q(10)$, which finishes the proof of the theorem. \square

Here is an explicit example. Consider $x = \frac{1}{7}$, then one easily checks that $\text{ord}_7(10) = 6$ and

$$10^6 - 1 = 7 \cdot 142857.$$

Thus

$$x = \frac{1}{7} = \frac{142857}{10^6 - 1} = \frac{142857}{10^6} + \frac{142857}{10^{12}} + \dots = 0.142857142857\dots$$

Example 7.74. (Moscow 1990) The decimal representation of a rational number A is purely periodic with period n . What is the longest possible length of the period of A^2 ?

Proof. Letting $A = \frac{a}{b}$, the hypothesis becomes $\text{ord}_b(10) = n$ and we need to find the maximal value of $\text{ord}_{b^2}(10)$. Write $10^n = 1 + kb$ and observe that by the binomial formula we have

$$10^{nb} = (1 + kb)^b = 1 + kb^2 + \dots \equiv 1 \pmod{b^2}.$$

Since $b^2 \mid 10^{bn} - 1$, it follows that $\text{ord}_{b^2}(10) \mid bn$, in particular $\text{ord}_{b^2}(10) \leq bn \leq n(10^n - 1)$. To see that this is the answer, it remains to prove that we can find A for which $\text{ord}_b(10) = n$ and $\text{ord}_{b^2}(10) = n(10^n - 1)$. Take $A = \frac{1}{10^n - 1}$, so $b = 10^n - 1$. Let $k = \text{ord}_{b^2}(10)$, then clearly $n \mid k$ and so $k = nc$ for some positive integer c . Moreover $(10^n - 1)^2 \mid 10^{nc} - 1$, thus $10^n - 1 \mid 1 + 10^n + \dots + 10^{n(c-1)}$, which yields $10^n - 1 \mid c$ and finally $\text{ord}_{b^2}(10) = n(10^n - 1)$. \square

Example 7.75. (USAMO 2013) Let m and n be positive integers. Prove that there is a positive integer c such that the numbers cm and cn have the same number of occurrences of each non-zero digit when written in base ten.

Proof. Start by choosing a positive integer k such that $10^k m - n$ can be written $10^k m - n = 2^x 5^y z$ with $x, y \geq 0$, z relatively prime to 10 and $z > \max(m, n)$. This is possible, since for $k > \max(v_2(n), v_5(n))$ we have $v_p(10^k m - n) = v_p(n)$

if $p \in \{2, 5\}$, thus $z \geq \frac{10^k m - n}{2^{v_2(n)} 5^{v_5(n)}}$ and this last quantity exceeds $\max(m, n)$ for k large enough.

Next, let b be the order of 10 modulo z and write $10^b - 1 = zc$ for some positive integer c . We claim that this c is a solution of the problem. First, observe that b is the number of digits in the period of $\frac{1}{z}$, and this period is the b -digit decimal representation of c (with possibly some extra zeroes added to the left of the usual decimal representation of c). Since $z > \max(m, n)$, the decimal expansions of $\frac{m}{z}$ and $\frac{n}{z}$ consist of repeated b -digit representations of cm and cn . Since

$$10^k \frac{m}{z} = \frac{n}{z} + 2^x 5^y,$$

the decimal expansion of $\frac{n}{z}$ is obtained from that of $\frac{m}{z}$ by shifting the decimal to the right k places and removing the integer part. It follows that the b -digit representations of cm and cn are cyclic shifts of one another, which shows that c is a solution of the problem. \square

Example 7.76. (IMO Shortlist 1999) a) Prove that there are infinitely many primes p such that the length of the period of $\frac{1}{p}$ is a multiple of 3.

b) If p is such a prime number, write $\frac{1}{p} = 0.a_1a_2\dots a_{3k}a_1a_2\dots a_{3k}\dots$. What is the maximal value of $\max_{1 \leq i \leq k} (a_i + a_{i+k} + a_{i+2k})$ over all such primes p ?

Proof. a) We need to ensure that the order of 10 modulo p is a multiple of 3. If this order is $3d$, then p divides $10^{2d} + 10^d + 1$, which suggests looking at divisors of $10^{2q} + 10^q + 1$, with q a prime (so that $3q$ has few divisors). More precisely, we will prove that for any prime q we can find a prime divisor $p = f(q)$ of $10^{2q} + 10^q + 1$ which does not divide $10^3 - 1$. Moreover, we will prove that the order of 10 modulo p is $3q$, in particular $q \rightarrow f(q)$ is injective, which will yield part a).

Note that $10^{2q} + 10^q + 1 \equiv 3 \pmod{9}$, thus if all prime divisors of $10^{2q} + 10^q + 1$ divide $10^3 - 1 = 9 \cdot 111 = 27 \cdot 37$, then $10^{2q} + 10^q + 1 = 3 \cdot 37^k$ for some positive integer k , which is impossible (take the equation mod 4). This proves the existence of p .

Next, let d be the order of 10 modulo p . Since p divides $10^{2q} + 10^q + 1$, it also divides $10^{3q} - 1$ and so d divides $3q$. If $d \neq 3q$, then $d = 1, 3$ or q . The first two cases are impossible by the choice of p . If $d = q$, then $10^q \equiv 1 \pmod{p}$

and since p divides $10^{2q} + 10^q + 1$, it follows that $p \mid 3$, a contradiction. Hence $d = 3q$ and we are done.

b) This part is trickier.

As we have already observed, we have $p \mid 10^{2k} + 10^k + 1$. Since

$$\frac{10^{3k} - 1}{p} = a_1 \cdot 10^{3k-1} + \dots + a_{3k},$$

we deduce that

$$10^k - 1 \mid a_1 \cdot 10^{3k-1} + \dots + a_{3k},$$

which can be rewritten (using that $10^{kj+r} \equiv 10^r \pmod{10^k - 1}$) as

$$10^k - 1 \mid b_1 \cdot 10^{k-1} + b_2 \cdot 10^{k-2} + \dots + b_k,$$

where $b_i = a_i + a_{i+k} + a_{i+2k}$. Note that $0 \leq b_i \leq 27$, thus

$$b_1 \cdot 10^{k-1} + b_2 \cdot 10^{k-2} + \dots + b_k \leq 27 \cdot \frac{10^k - 1}{9} = 3(10^k - 1).$$

Moreover, we have equality if and only if $a_1 = \dots = a_{3k} = 9$, which is impossible (it would force $p = 1$). Thus $b_1 \cdot 10^{k-1} + b_2 \cdot 10^{k-2} + \dots + b_k$ is a multiple of $10^k - 1$ smaller than $3(10^k - 1)$, so it cannot exceed $2(10^k - 1)$. In particular

$$b_1 \leq \frac{2(10^k - 1)}{10^{k-1}} < 20.$$

On the other hand, since $10^k - 1 \mid 10(b_1 \cdot 10^{k-1} + b_2 \cdot 10^{k-2} + \dots + b_k)$, we also obtain $10^k - 1 \mid b_2 \cdot 10^{k-1} + \dots + 10b_k + b_1$ and so the previous argument yields $b_2 < 20$. Continuing like this we obtain $b_3, \dots, b_k < 20$, thus

$$\max_{1 \leq i \leq k} (a_i + a_{i+k} + a_{i+2k}) \leq 19.$$

We conclude observing that for $p = 7$ the maximum is attained, since then $a_1 = 1$, $a_2 = 4$, $a_3 = 2$, $a_4 = 8$, $a_5 = 5$ and $a_6 = 7$, and $4 + 8 + 7 = 19$. \square

7.3.2 Practicing the notion of order modulo n

In this section we illustrate the previous theoretical results with some concrete, but more challenging examples. The result established in the next problem is extremely useful in practice. Roughly, it says that if a, b are integers, then the prime factors of $a^p - b^p$ (p being a prime) are of a rather special form.

Example 7.77. Let a and b be different integers and let p be a prime.

a) Prove that any prime factor q of $a^p - b^p$ is either a divisor of $\gcd(a, b) \cdot (a - b)$ or of the form $1 + kp$, with $k \geq 1$.

b) Suppose that $\gcd(a, b) = 1$. Prove that any prime factor q of $\frac{a^p - b^p}{a - b}$ is either equal to p or of the form $1 + kp$ with $k \geq 1$.

Proof. a) If $q \mid a$, then clearly $q \mid b$ and so $q \mid \gcd(a, b)$. Assume now that q does not divide a , then it does not divide b either (since $q \mid a^p - b^p$). Let c be an integer such that $ca \equiv 1 \pmod{q}$, then $q \mid (ca)^p - (cb)^p$, thus $q \mid (cb)^p - 1$. If $d = \text{ord}_q(cb)$, then $d \mid p$ and $d \mid \varphi(q) = q - 1$. If $d = 1$ then $q \mid cb - 1$ and then $q \mid a - b$ since $ac \equiv 1 \pmod{q}$. If $d = p$ then $p \mid q - 1$ and we are done again.

b) Since $q \mid \frac{a^p - b^p}{a - b}$ we have $q \mid a^p - b^p$ and so, by part a) and the hypothesis, $q \mid a - b$ or $q \equiv 1 \pmod{p}$. If $q \equiv 1 \pmod{p}$ we are done, so assume that $q \mid a - b$. We also know that $q \mid a^{p-1} + a^{p-2}b + \dots + b^{p-1}$, thus $q \mid pa^{p-1}$ and $q \mid pb^{p-1}$. Since $\gcd(a^{p-1}, b^{p-1}) = 1$, we conclude that $q \mid p$ and finally $q = p$. The result follows. \square

Remark 7.78. In part b) if we assume that $p, q > 2$ then $q \equiv 1 \pmod{2p}$ and so $q \geq 2p + 1$.

A very similar and also very useful result is the following:

Example 7.79. Let a and b be relatively prime integers and let n be a positive integer. Prove that any odd positive divisor of $a^{2^n} + b^{2^n}$ is congruent to 1 modulo 2^{n+1} .

Proof. It suffices to prove that any odd prime divisor p of $a^{2^n} + b^{2^n}$ is congruent to 1 modulo 2^{n+1} . Note that p does not divide ab , since $\gcd(a, b) = 1$. Let c be an integer such that $bc \equiv 1 \pmod{p}$, then $p \mid (ac)^{2^n} + 1$. Then the order k of ac modulo p divides 2^{n+1} , since $p \mid (ac)^{2^{n+1}} - 1$, and does not divide 2^n , since

otherwise we would have $p \mid (ac)^{2^n} - 1$ and $p \mid (ac)^{2^n} + 1 - ((ac)^{2^n} - 1) = 2$, a contradiction. Thus $k = 2^{n+1}$ and since $k \mid p - 1$ we are done. \square

The next four examples are illustrations of the result established in the previous example.

Example 7.80. (Kvant M 1476) Find all primes p and q such that

$$pq \mid (2^p + 1)(2^q + 1).$$

Proof. If $p \mid 2^p + 1$ then Fermat's little theorem yields $p \mid 3$ and then $p = 3$. Thus if $p \mid 2^p + 1$ then $p = 3$ and $q \mid 3(2^q + 1)$. Using again Fermat's little theorem we obtain $q \mid 9$ and then $q = 3$, giving the solution $(p, q) = (3, 3)$. On the other hand, if $(p, q) \neq (3, 3)$, the previous discussion shows that we must have $p, q \neq 3$, $p \neq q$, $p \mid 2^q + 1$ and $q \mid 2^p + 1$. We will prove that this is impossible. Since $p \neq 3$ and $p \mid 2^q + 1$, we have $p \mid \frac{(-2)^q - 1}{-2 - 1}$ and example 7.77 yields $p \equiv 1 \pmod{q}$, in particular $p > q$. By symmetry we also obtain $q > p$, a contradiction. Thus $(p, q) = (3, 3)$ is the only solution of the problem. \square

Example 7.81. (IMO Shortlist 2006) Find all integer solutions of the equation

$$\frac{x^7 - 1}{x - 1} = y^5 - 1.$$

Proof. We will prove that the equation has no solutions, by using twice the following special case of example 7.77: if p is a prime and x is an integer then any prime factor q of $\frac{x^p - 1}{x - 1}$ is congruent to 0 or 1 modulo p . It follows that for any positive divisor d of $\frac{x^p - 1}{x - 1}$ we have $d \equiv 0, 1 \pmod{p}$.

Note that $\frac{x^7 - 1}{x - 1} > 0$ for any $x \neq 1$, since $x - 1$ and $x^7 - 1$ have the same sign, thus $y > 1$. The previous discussion shows that $y - 1$ and $z := y^4 + y^3 + y^2 + y + 1$ are each congruent to 0 or 1 modulo 7. If $y - 1 \equiv 0 \pmod{7}$, then $z \equiv 5 \pmod{7}$, a contradiction. If $y - 1 \equiv 1 \pmod{7}$, then $z \equiv 2^4 + 2^3 + 2^2 + 2 + 1 \equiv 3 \pmod{7}$, again a contradiction. Hence the equation has no solution. \square

Example 7.82. Find all integers $a, n > 1$ such that n and $a^n + 1$ have the same set of prime divisors.

Proof. Let p be the largest prime divisor of n . If $p = 2$ then n is a power of 2, as well as $a^n + 1 = (a^{\frac{n}{2}})^2 + 1$. Since 4 cannot divide $x^2 + 1$ for any integer x we deduce that $a^n + 1 = 2$ and $a = 1$, a contradiction. Thus $p > 2$. Let $b = a^{\frac{n}{p}}$ and consider

$$A = \frac{b^p + 1}{b + 1} = \frac{(-b)^p - 1}{(-b) - 1} = \frac{a^n + 1}{a^{\frac{n}{p}} + 1}.$$

Any prime factor q of A is either equal to p or congruent to 1 modulo p by example 7.77. On the other hand, q also divides $a^n + 1$, thus $q \mid n$ and then $q \leq p$. It follows that $q = p$ and so A is a power of p . Moreover, $p \mid b^p + 1$ and so $p \mid b + 1$ by Fermat's little theorem. Using the lifting the exponent lemma we obtain $v_p(A) = 1$ and so the only possibility is $A = p$, that is $b^p + 1 = p(b + 1)$. Arguing as in the solution of example 6.29 (this is a simple argument based on inequalities) yields $b = 2$ and $p = 3$, then $a^{\frac{n}{p}} = 2$ and $a = 2$, $n = 3$. Thus the only solution of the problem is $(a, n) = (2, 3)$. \square

Example 7.83. (IMO Shortlist 2005) Find all positive integers n for which there is a unique integer $a \in \{0, 1, \dots, n! - 1\}$ satisfying $n! \mid a^n + 1$.

Proof. It is not difficult to see that $n = 2$ and $n = 3$ are solutions, so assume that $n > 3$. If $n! \mid a^n + 1$ then $4 \mid a^n + 1$ and so n must be odd. Hence $a = n! - 1$ satisfies $n! \mid a^n + 1$, which shows that n is a solution of the problem if and only if $b^n + 1$ is not a multiple of $n!$ for any $b \in \{0, 1, \dots, n! - 2\}$.

Suppose first that n is a prime and that $b \in \{0, 1, \dots, n! - 2\}$ satisfies $n! \mid b^n + 1$. Then $n \mid b^n + 1$ and Fermat's little theorem gives $n \mid b + 1$. On the other hand, choose any prime $q < n$ and let $k = v_q((n - 1)!)$, then $q^k \mid (b + 1) \cdot \frac{b^n + 1}{b + 1}$. Since $q < n$, example 7.77 shows that q cannot divide $\frac{b^n + 1}{b + 1}$ and so $q^k \mid b + 1$. It follows that $(n - 1)! \mid b + 1$, which combined with $n \mid b + 1$ and $\gcd(n, (n - 1)!) = 1$ gives $n! \mid b + 1$, a contradiction. Thus all prime numbers are solutions of the problem.

Suppose next that n is composite and let p be the smallest prime factor of n . We will prove that $b = \frac{n!}{p} - 1 \in \{0, 1, \dots, n! - 2\}$ satisfies $n! \mid b^n + 1$ and so n is not a solution. But

$$b^n + 1 = (b + 1)(b^{n-1} - b^{n-2} + \dots + 1) = \frac{n!}{p} \cdot (b^{n-1} - b^{n-2} + \dots + 1)$$

and so it suffices to prove that $p \mid b^{n-1} - b^{n-2} + \dots + 1$. Since p is the smallest prime factor of n and n is composite, we have $p^2 \leq n$ and so $p^2 \mid n!$, thus $b \equiv -1 \pmod{p}$ and $b^{n-1} - b^{n-2} + \dots + 1 \equiv 1 + 1 + \dots + 1 = n \equiv 0 \pmod{p}$, as desired. Thus the solutions of the problem are exactly the prime numbers. \square

Example 7.84. (Komal) Let $n \geq 1$ and a be integers such that $n \mid a^n - 1$. Prove that $a + 1, a^2 + 2, a^3 + 3, \dots, a^n + n$ form a complete residue system modulo n .

Proof. We will prove this by strong induction on n , the case $n = 1$ being clear. Assume that the result holds up to $n - 1$ and let us prove it for n . Note that $\gcd(a, n) = 1$ since $n \mid a^n - 1$, thus we can set $d = \text{ord}_n(a)$, and we have $d \mid \varphi(n)$, in particular $d < n$. Moreover, since $a^n \equiv 1 \pmod{n}$ we have $d \mid n$, which yields $a^d \equiv 1 \pmod{d}$ (since $a^d \equiv 1 \pmod{n}$ and $d \mid n$). Thus a and d satisfy the same hypotheses as a and n , and moreover $d < n$. The inductive hypothesis shows that $(a^i + i)_{1 \leq i \leq d}$ is a complete residue system modulo d .

Assume next that $a^i + i \equiv a^j + j \pmod{n}$ for some integers $i, j \geq 0$, then $a^i + i \equiv a^j + j \pmod{d}$ (since $d \mid n$) and by the previous discussion $i \equiv j \pmod{d}$. But then $a^i \equiv a^j \pmod{n}$ (since $a^d \equiv 1 \pmod{n}$), which combined with the congruence $a^i + i \equiv a^j + j \pmod{n}$ yields $i \equiv j \pmod{n}$. The result follows. \square

Example 7.85. (India 2014) Let p be an odd prime and let k be an odd positive integer. Prove that $pk + 1$ does not divide $p^p - 1$.

Proof. Suppose that this is not the case and let k be the smallest odd positive integer for which $pk + 1 \mid p^p - 1$. The order of p modulo $pk + 1$ divides p and cannot be 1 (since $pk + 1$ does not divide $p - 1$), thus it must be p , which shows that $p \mid \varphi(pk + 1)$. Since $\gcd(p, pk + 1) = 1$, it follows that there is a prime $q \mid pk + 1$ such that $p \mid q - 1$. In particular we have $q > 2$ and so $2p \mid q - 1$. Write $pk + 1 = q^s m$ with $s \geq 1$ and $m \geq 1$ not divisible by q . Taking the equation $pk + 1 = q^s m$ modulo $2p$ and using that k is odd and $q \equiv 1 \pmod{2p}$ yields $m \equiv 1 + p \pmod{2p}$, thus $m = 1 + up$ for some positive odd integer u . Since $m < pk + 1$ we have $u < k$ and since $m \mid pk + 1$ we also have $1 + up \mid p^p - 1$. This contradicts the minimality of k and finishes the proof. \square

We end this section with some more challenging problems.

Example 7.86. (Romania TST 2009) Prove that there are infinitely many pairs of distinct prime numbers (p, q) such that $p \mid 2^{q-1} - 1$ and $q \mid 2^{p-1} - 1$.

Proof. Let $F_n = 2^{2^n} + 1$ be the n th Fermat number. For each $n > 1$ let p_n be a prime factor of F_n and let q_n be a prime factor of F_{n+1} . Then p_2, p_3, \dots are pairwise distinct and $p_n \neq q_n$ for all n , since the Fermat numbers are pairwise relatively prime (see example 3.12). Moreover by example 7.61 we have $p_n \equiv 1 \pmod{2^{n+2}}$ and $q_n \equiv 1 \pmod{2^{n+3}}$. Thus $p_n \mid 2^{2^n} + 1 \mid 2^{2^{n+1}} - 1 \mid 2^{q_n-1} - 1$ and $q_n \mid 2^{2^{n+1}} + 1 \mid 2^{2^{n+2}} - 1 \mid 2^{p_n-1} - 1$. Thus (p_n, q_n) is a solution of the problem for all $n > 1$. \square

Example 7.87. (Russia 2009) Let x and y be integers such that $2 \leq x, y \leq 100$. Prove that there exists a positive integer n such that $x^{2^n} + y^{2^n}$ is a composite number.

Proof. The result is clear when $x = y$ (take $n = 1$), so assume that $x \neq y$. We first prove that $257 \mid x^{2^n} + y^{2^n}$ for some $n \geq 1$. Since 257 is a prime and y is not divisible by 257 there is q such that $x \equiv qy \pmod{257}$. Note that q is not congruent $0, \pm 1 \pmod{257}$ thanks to the hypothesis $2 \leq x, y \leq 100$ and $x \neq y$. Let $d = \text{ord}_{257}(q)$, then $d \mid 256 = 2^8$ and so $d = 2^k$ for some k . Since 257 does not divide $q \pm 1$, we have $k \geq 2$. Moreover, since $257 \mid q^{2^k} - 1$ and 257 does not divide $q^{2^{k-1}} - 1$, we have $257 \mid q^{2^{k-1}} + 1$. Finally, since $x \equiv qy \pmod{257}$, it follows that $257 \mid x^{2^{k-1}} + y^{2^{k-1}}$ and the claim is proved (take $n = k - 1 \geq 1$).

Suppose now that $x^{2^n} + y^{2^n}$ is a prime, then necessarily $x^{2^n} + y^{2^n} = 257$. Letting $a = x^{2^{n-1}}$ and $b = y^{2^{n-1}}$, we obtain $a^2 + b^2 = 257$ and $a, b > 1$ (since $x, y > 1$). One easily checks by hand that this is impossible (the general result is that a prime $p \equiv 1 \pmod{4}$ can be written in a unique way as a sum of two squares and in this case $257 = 16^2 + 1^2$ is that way), which shows that $x^{2^n} + y^{2^n}$ is a composite number. \square

Example 7.88. (AMME 2948) Let x, y be relatively prime integers greater than 1. Prove that $v_p(x^{p-1} - y^{p-1})$ is odd for infinitely many primes p .

Proof. If $k > 2$ is an integer, by theorem 3.51 and the remark following it $x^{2^{k-1}} + y^{2^{k-1}}$ is neither a perfect square nor twice a perfect square. Thus we can find an odd prime p_k such that $v_{p_k}(x^{2^{k-1}} + y^{2^{k-1}})$ is odd. Since $\gcd(x, y) = 1$, p_k cannot divide xy . Since it divides $x^{2^{k-1}} + y^{2^{k-1}}$, example 7.79 shows that 2^k divides $p_k - 1$. The lifting the exponent lemma gives

$$v_{p_k}(x^{p_k-1} - y^{p_k-1}) = v_{p_k}(x^{2^k} - y^{2^k}) + v_{p_k}\left(\frac{p_k - 1}{2^k}\right) = v_{p_k}(x^{2^{k-1}} + y^{2^{k-1}}),$$

and the last quantity is odd by the choice of p_k . The result follows by taking successively $k = 3, 4, \dots$ and observing that $p_k \geq 1 + 2^k$. \square

Example 7.89. (China TST 2005) Prove that for any $n > 2$ the number $2^{2^n} + 1$ has a prime factor greater than $(n + 1) \cdot 2^{n+2}$.

Proof. The result is clear for $n = 3$ (note that $2^8 + 1$ is a prime), so assume that $n \geq 4$. Consider the prime factorization

$$2^{2^n} + 1 = p_1^{\alpha_1} \dots p_k^{\alpha_k},$$

with $p_1 < \dots < p_k$. By example 7.61 there are positive integers q_1, \dots, q_k such that $p_i = 1 + 2^{n+2}q_i$. Since $2^n \geq 2n + 4$ (as $n \geq 4$) and $p_i^{\alpha_i} \equiv 1 + 2^{n+2}\alpha_i q_i \pmod{2^{2n+4}}$, we obtain

$$1 \equiv 2^{2^n} + 1 \equiv \prod_{i=1}^k (1 + 2^{n+2}\alpha_i q_i) \equiv 1 + 2^{n+2} \sum_{i=1}^k \alpha_i q_i \pmod{2^{2n+4}},$$

thus

$$\alpha_1 q_1 + \dots + \alpha_k q_k \geq 2^{n+2}.$$

Assuming that $\max_i(q_i) \leq n$, we obtain $\alpha_1 + \dots + \alpha_k \geq \frac{2^{n+2}}{n}$, which gives

$$1 + 2^{2^n} = \prod_{i=1}^k (1 + 2^{n+2}q_i)^{\alpha_i} > \prod_{i=1}^k (2^{n+2})^{\alpha_i} \geq (2^{n+2})^{\frac{2^{n+2}}{n}} = 2^{\frac{n+2}{n} \cdot 2^{n+2}}$$

and so $1 + 2^{2^n} > 2^{2^{n+2}}$, a contradiction. Thus $\max_i(q_i) \geq n + 1$ and so $\max_i(p_i) > (n + 1)2^{n+2}$, as desired. \square

Example 7.90. (Iran 2011) Let $k \geq 7$ be an integer. Find the number of pairs (x, y) such that $0 \leq x, y < 2^k$ and

$$73^{73^x} \equiv 9^{9^y} \pmod{2^k}.$$

Proof. We start by finding the possible remainders of the numbers $1, 9, 9^2, \dots$ when divided by 2^N , for a given integer $N \geq 4$. We easily obtain (using proposition 7.68 or, better, by a direct computation) that the order of $9 \pmod{2^N}$ is 2^{N-3} . Thus there are precisely 2^{N-3} distinct residues modulo 2^N among those of $1, 9, 9^2, \dots$. As each of these residues is of the form $8k + 1$ and since there are 2^{N-3} such residues, it follows that the remainders of $1, 9, 9^2, \dots$ are exactly the residues mod 2^N of the form $8k + 1$.

Since $73 \equiv 1 \pmod{8}$, the previous paragraph gives the existence of $u \geq 1$ such that $73 \equiv 9^u \pmod{2^k}$. Since $73 \equiv 9 \pmod{64}$, it follows that $9^{u-1} \equiv 1 \pmod{2^6}$ and the previous paragraph (with $N = 6$) yields $u \equiv 1 \pmod{8}$. Since the order of 9 modulo 2^k is 2^{k-3} , the congruence $73^{73^x} \equiv 9^{9^y} \pmod{2^k}$ is equivalent to $u9^{ux} \equiv 9^y \pmod{2^{k-3}}$. We need to find the number of solutions of this congruence with $x, y \in \{0, 1, \dots, 2^k - 1\}$. Fix $x \in \{0, 1, \dots, 2^k - 1\}$. Then $u9^{ux} \equiv 1 \pmod{8}$, hence by the first paragraph we can find v such that $u9^{ux} \equiv 9^v \pmod{2^{k-3}}$. Now $9^y \equiv 9^v \pmod{2^{k-3}}$ if and only if $y \equiv v \pmod{2^{k-6}}$. There are precisely 2^6 such numbers $y \in \{0, 1, \dots, 2^k - 1\}$. Thus for each x the corresponding congruence has 2^6 solutions y , and so the total number of solutions is 2^{k+6} . \square

Example 7.91. (Iran TST 2009) Prove that for all positive integers n we have

$$3^{\frac{5^{2^n}-1}{2^{n+2}}} \equiv (-5)^{\frac{3^{2^n}-1}{2^{n+2}}} \pmod{2^{n+4}}.$$

Proof. Denote for simplicity $5^{2^n} - 1 = b$ and $3^{2^n} - 1 = c$. One easily checks, using either the lifting the exponent lemma or the formula

$$x^{2^n} - 1 = (x - 1)(x + 1)(x^2 + 1) \dots (x^{2^{n-1}} + 1),$$

that $v_2(b) = v_2(c) = n + 2$, thus $\frac{b}{2^{n+2}}$ and $\frac{c}{2^{n+2}}$ are odd integers and the congruence can also be written as

$$(-3)^{\frac{b}{2^{n+2}}} \equiv 5^{\frac{c}{2^{n+2}}} \pmod{2^{n+4}}.$$

Next, by example 7.60 there is $a \geq 1$ such that $-3 \equiv 5^a \pmod{2^{n+4}}$. The previous congruence becomes

$$5^{\frac{ab}{2^{n+2}}} \equiv 5^{\frac{c}{2^{n+2}}} \pmod{2^{n+4}}.$$

Since the order of 5 modulo 2^{n+4} is 2^{n+2} (see example 7.60), this last congruence is equivalent to

$$\frac{ab}{2^{n+2}} \equiv \frac{c}{2^{n+2}} \pmod{2^{n+2}}$$

or $ab \equiv c \pmod{2^{2n+4}}$.

Next, observe that if x, y are odd integers with $x \equiv y \pmod{2^k}$ for some $k \geq 1$, then $x^{2^m} \equiv y^{2^m} \pmod{2^{m+k}}$ for all $m \geq 1$. This follows immediately by induction, or using the formula

$$x^{2^m} - y^{2^m} = (x - y)(x + y)(x^2 + y^2) \dots (x^{2^{m-1}} + y^{2^{m-1}}).$$

Since $-3 \equiv 5^a \pmod{2^{n+4}}$, we deduce that $3^{2^n} \equiv 5^{a \cdot 2^n} \pmod{2^{2n+4}}$. Hence

$$1 + c \equiv (1 + b)^a = 1 + ab + \binom{a}{2} b^2 + \dots \pmod{2^{2n+4}}.$$

Since $v_2(b) = n + 2$, the last congruence is equivalent to $c \equiv ab \pmod{2^{2n+4}}$, which is exactly what we needed. \square

Example 7.92. (China TST 2004) Prove that for any integer $m > 1$ there is a prime number p which does not divide $n^m - m$ for any integer n .

Proof. Choose a prime factor q of m . We will prove in the next paragraph that we can find a prime p such that $p \mid m^q - 1$, p does not divide $m - 1$ and finally $\gcd(p - 1, qm) \mid m$. We claim that such p is a solution of the problem. Indeed, assuming that $p \mid n^m - m$ for some n , we obtain $n^{mq} \equiv m^q \equiv 1 \pmod{p}$, so $d := \text{ord}_p(n)$ satisfies $d \mid mq$. Since $d \mid p - 1$, we have $d \mid \gcd(mq, p - 1) \mid m$ and so $p \mid n^m - 1$. Since $p \mid n^m - m$, we conclude that $p \mid m - 1$, contradicting the choice of p .

We prove now the existence of p . Letting $k = v_q(m)$, the number

$$A = \frac{m^q - 1}{m - 1} = 1 + m + m^2 + \dots + m^{q-1}$$

is congruent to $1+m$ modulo q^{k+1} and so it is not congruent to 1 modulo q^{k+1} . It follows that A has a prime factor p which is not congruent to 1 modulo q^{k+1} . Then clearly $p \mid m^q - 1$ and $\gcd(p-1, mq) \mid m$. We cannot have $p \mid m-1$, since otherwise $p = q$ (as $p \mid A = 1 + m + \dots + m^{q-1}$ and $p \mid m-1$ force $p \mid q$) and $q \mid m-1$, a contradiction with $q \mid m$. Thus p satisfies all desired conditions and the problem is solved. \square

Remark 7.93. The case when m is a prime was one of the problems given at IMO 2003.

7.3.3 Primitive roots modulo n

We have already seen that for any positive integer n and any integer a relatively prime to n the order modulo n of a divides $\varphi(n)$ and in particular it cannot exceed $\varphi(n)$. We are interested in this section in characterizing those n for which this bound is attained, i.e. for which there is a such that $\gcd(a, n) = 1$ and $\text{ord}_n(a) = \varphi(n)$. Let us give a name to such numbers a .

Definition 7.94. Let n be a positive integer. An integer a is called a primitive root modulo n if $\gcd(a, n) = 1$ and $\text{ord}_n(a) = \varphi(n)$.

It is clear that if a is a primitive root modulo n and if $b \equiv a \pmod{n}$, then b is also a primitive root modulo n . Note that an integer a relatively prime to n is a primitive root modulo n if and only if $1, a, \dots, a^{\varphi(n)-1}$ give pairwise distinct remainders modulo n . This yields the following useful observation.

Proposition 7.95. Let a be an integer relatively prime to a positive integer n . The following statements are equivalent:

- a) a is a primitive root modulo n ;
- b) $1, a, a^2, \dots, a^{\varphi(n)-1}$ is a reduced residue system modulo n ;
- c) For any integer x relatively prime to n there is a positive integer k such that $x \equiv a^k \pmod{n}$.

Let us give a few simple examples: the primitive roots modulo 2 are the odd integers, the primitive roots modulo 3 (respectively 4) are the integers of the form $3k+2$ (respectively $4k+3$). Similarly, the primitive roots modulo 5

are integers of the form $5k + 2$ or $5k + 3$ and the primitive roots modulo 6 are integers of the form $6k + 5$.

The next proposition gives a useful criterion for proving that an integer a is a primitive root mod n .

Proposition 7.96. *Let $n > 1$ be an integer and let a be an integer such that $\gcd(a, n) = 1$. Then a is a primitive root mod n if and only if n does not divide $a^{\frac{\varphi(n)}{q}} - 1$ for all primes $q | \varphi(n)$.*

Proof. If a is a primitive root mod n , then n does not divide $a^{\frac{\varphi(n)}{q}} - 1$ since otherwise $\varphi(n) = \text{ord}_n(a)$ would divide $\frac{\varphi(n)}{q}$.

Conversely, suppose that n does not divide $a^{\frac{\varphi(n)}{q}} - 1$ for all primes $q | \varphi(n)$, and let $d = \text{ord}_n(a)$. Then $d | \varphi(n)$ and by assumption d does not divide $\frac{\varphi(n)}{q}$ for any prime factor q of $\varphi(n)$. It follows that $\frac{\varphi(n)}{d}$ is a divisor of $\varphi(n)$ which is not divisible by any prime factor of $\varphi(n)$, thus $\frac{\varphi(n)}{d} = 1$ and a is a primitive root mod n . \square

We illustrate the previous proposition with a few concrete examples, some of which use intensively results about quadratic residues discussed in chapter 4.

Example 7.97. Prove that 2 is a primitive root modulo 29 and solve the congruence

$$1 + x + \dots + x^6 \equiv 0 \pmod{29}.$$

Proof. By proposition 7.96, it suffices to check that 2^{14} and 2^4 are not congruent to 1 modulo 29. This is clear for 2^4 , and follows for 2^{14} from

$$2^{14} \equiv (2^5)^2 \cdot 2^4 \equiv 3^2 \cdot 16 = 3 \cdot 48 \equiv -30 \equiv -1 \pmod{29}.$$

Thus 2 is a primitive root modulo 29 (one could also observe that $29 \equiv 5 \pmod{8}$, hence $\left(\frac{2}{29}\right) = -1$ and $2^{14} \equiv -1 \pmod{29}$).

Suppose that x is not congruent to 1 modulo 29, then $1 + x + \dots + x^6 \equiv 0 \pmod{29}$ if and only if $x^7 \equiv 1 \pmod{29}$. Write $x \equiv 2^k \pmod{29}$ for some $0 \leq k \leq 27$, which is possible since 2 is a primitive root modulo 29. Then $x^7 \equiv 1 \pmod{29}$ if and only if $28 | 7k$, that is $4 | k$. We deduce that the solutions of the congruence are 2^{4k} for $1 \leq k \leq 6$. \square

Example 7.98. (Putnam 1994) For a nonnegative integer a let

$$n_a = 101a - 100 \cdot 2^a.$$

Prove that if $a, b, c, d \in \{0, 1, \dots, 99\}$ satisfy $n_a + n_b \equiv n_c + n_d \pmod{10100}$, then $\{a, b\} = \{c, d\}$.

Proof. The congruence $n_a + n_b \equiv n_c + n_d \pmod{10100}$ is equivalent to the simultaneous congruences $a + b \equiv c + d \pmod{100}$ and $2^a + 2^b \equiv 2^c + 2^d \pmod{101}$. Since 101 is a prime number, Fermat's little theorem combined with $a + b \equiv c + d \pmod{100}$ yield $2^a \cdot 2^b \equiv 2^c \cdot 2^d \pmod{101}$. It follows that

$$(X - 2^a)(X - 2^b) \equiv (X - 2^c)(X - 2^d) \pmod{101}$$

and so $(2^a - 2^c)(2^a - 2^d) \equiv 0 \pmod{101}$. By symmetry, we may assume that $2^a \equiv 2^c \pmod{101}$, thus $\text{ord}_{101}(2) \mid a - c$. We will prove below that $\text{ord}_{101}(2) = 100$, which yields $a = c$ and then $b = d$.

It remains to prove that 2 is a primitive root modulo 101. By proposition 7.96, it suffices to prove that $2^{20} - 1$ and $2^{50} - 1$ are not multiples of 101. For $2^{20} - 1$ we observe that

$$2^{20} = (2^{10})^2 \equiv 14^2 \equiv 95 \pmod{101}.$$

For $2^{50} - 1$ one can use a similar computation, or, more elegantly, use Euler's criterion (theorem 5.99) and the fact that $\left(\frac{2}{101}\right) = -1$ (use theorem 5.125 and the congruence $101 \equiv 5 \pmod{8}$). \square

Example 7.99. Let $p > 3$ be a Fermat prime, i.e. of the form $2^n + 1$ for some integer $n > 1$. Prove that 3 is a primitive root mod p .

Proof. Since $\varphi(p) = p - 1 = 2^n$, by proposition 7.96 it suffices to prove that $3^{\frac{p-1}{2}} - 1$ is not divisible by p , which is equivalent by Euler's criterion (theorem 5.99) to $\left(\frac{3}{p}\right) = -1$. Using the quadratic reciprocity law (theorem 5.124) we obtain

$$\left(\frac{3}{p}\right) = (-1)^{\frac{p-1}{2}} \cdot \left(\frac{p}{3}\right) = -1,$$

the last equality being a consequence of the fact that $p \equiv 1 \pmod{4}$ and $p \equiv 2 \pmod{3}$. The result follows. \square

Example 7.100. Let $q \equiv 1 \pmod{4}$ be a prime such that $p = 2q + 1$ is also prime. Prove that 2 is a primitive root modulo p .

Proof. Again, by proposition 7.96 it suffices to prove that $2^{\frac{p-1}{2}} - 1$ and $2^{\frac{p-1}{q}} - 1$ are not divisible by p . This is clear for $2^{\frac{p-1}{q}} - 1 = 3$, so it remains to prove that $\left(\frac{2}{p}\right) = -1$ (by Euler's criterion, theorem 5.99). This follows from theorem 5.125 and the fact that $p \equiv 3 \pmod{8}$. \square

Remark 7.101. A famous conjecture of Artin implies the existence of infinitely many primes p for which 2 is a primitive root modulo p . The previous example shows that this would follow from the existence of infinitely many primes $q \equiv 1 \pmod{4}$ for which $p = 2q + 1$ is also a prime.

A natural question is whether for any positive integer n there are primitive roots modulo n . The answer is unfortunately negative: since any odd integer a satisfies $a^2 \equiv 1 \pmod{8}$, the order of any odd integer modulo 8 is 1 or 2, so there are no primitive roots modulo 8. Similarly one easily checks that there are no primitive roots modulo 2^n for $n > 2$. More precisely we have the following result.

Proposition 7.102. *Let n be a positive integer for which there are primitive roots modulo n . Then $n = 1, 2, 4, p^k$ or $2p^k$ for some odd prime p and some positive integer k .*

Proof. Suppose that n has primitive roots modulo n and is not of the form indicated in the proposition. Note that n is not a power of 2 greater than 4, by the discussion preceding the proposition. It is then immediate to see (by considering the prime factorization of n) that we can write $n = ab$ for two relatively prime numbers $a, b > 2$. Since $a, b > 2$, the numbers $\varphi(a)$ and $\varphi(b)$ are even, thus for all integers x relatively prime to n Euler's theorem yields

$$x^{\frac{\varphi(n)}{2}} = x^{\varphi(a) \cdot \frac{\varphi(b)}{2}} \equiv 1 \pmod{a}$$

and similarly $x^{\frac{\varphi(n)}{2}} \equiv 1 \pmod{b}$. Since $\gcd(a, b) = 1$, we infer that $x^{\frac{\varphi(n)}{2}} \equiv 1 \pmod{n}$ and so $\text{ord}_n(x) \mid \frac{\varphi(n)}{2}$ for any x relatively prime to n . It follows that there are no primitive roots modulo n , a contradiction. \square

A remarkable theorem due to Gauss states that the converse of the previous result holds, giving a complete description of all positive integers n for which there are primitive roots modulo n .

Theorem 7.103. (Gauss) *Let n be a positive integer. The following statements are equivalent:*

- a) *There are primitive roots modulo n .*
- b) *n is equal to $1, 2, 4, p^k$ or $2p^k$ for some odd prime p and some $k \geq 1$.*

We have already established one implication. The other implication lies deeper and we will establish it in a series of steps, each of which is interesting in its own right. The most delicate part is establishing the existence of primitive roots modulo odd primes, a task to which we focus our attention.

Theorem 7.104. *Let p be an odd prime. For any positive divisor d of $\varphi(p) = p-1$ there are exactly $\varphi(d)$ numbers $n \in \{1, 2, \dots, p-1\}$ such that $\text{ord}_p(n) = d$. In particular, there are $\varphi(p-1) \geq 1$ primitive roots modulo p .*

Proof. Let $f(d)$ be the number of integers $n \in \{1, 2, \dots, p-1\}$ with $\text{ord}_p(n) = d$. We will prove below that $f(d) \leq \varphi(d)$ for all $d \mid p-1$. Assuming this, we obtain

$$\sum_{d \mid p-1} f(d) \leq \sum_{d \mid p-1} \varphi(d) = p-1,$$

the last equality being a consequence of Gauss' theorem 4.112. Since $\text{ord}_p(n) \mid p-1$ for all $n \in \{1, 2, \dots, p-1\}$, we clearly have $\sum_{d \mid p-1} f(d) = p-1$. It follows that all the inequalities $f(d) \leq \varphi(d)$ must be equalities and the theorem is proved.

We still need to prove that $f(d) \leq \varphi(d)$. This is clear if $f(d) = 0$, so assume that there is $n \in \{1, 2, \dots, p-1\}$ with $\text{ord}_p(n) = d$. By Lagrange's theorem 5.69 and the fact that n, n^2, \dots, n^d are pairwise distinct solutions of the congruence $x^d \equiv 1 \pmod{p}$, it follows that all solutions of this congruence are given by the remainders mod p of n, n^2, \dots, n^d . Hence if $m \in \{1, 2, \dots, p-1\}$ has order d modulo p , then $m \equiv n^j \pmod{p}$ for some $1 \leq j \leq d$. Since $\text{ord}_p(m) = d$, proposition 7.66 gives $\gcd(j, d) = 1$, which proves that $f(d) \leq \varphi(d)$: all numbers $m \in \{1, 2, \dots, p-1\}$ with order d modulo p must be among the remainders modulo p of the numbers n^j with j a totative of d . The theorem is therefore proved. \square

Remark 7.105. 1) The most difficult part of the proof of theorem 7.104 is the existence of a primitive root modulo p . Indeed, if a is a primitive root modulo p , then any $n \in \{1, 2, \dots, p-1\}$ is congruent to a^k for some $0 \leq k \leq p-2$, and proposition 7.66 shows that $\text{ord}_p(n) = d$ if and only if $\frac{p-1}{\gcd(p-1, k)} = d$, i.e. $k = \frac{p-1}{d} \cdot e$, with e a totative of d . Thus there are $\varphi(d)$ such integers n .

2) Here is a slightly different, but quite nice way of proving theorem 7.104. Let $f(d)$ be the number of integers $n \in \{1, 2, \dots, p-1\}$ with $\text{ord}_p(n) = d$. We claim that for any $d \mid p-1$

$$\sum_{e \mid d} f(e) = d.$$

A number $x \in \{1, 2, \dots, p-1\}$ satisfies $x^d \equiv 1 \pmod{p}$ if and only if $e := \text{ord}_p(x)$ is a divisor of d , thus the left-hand side is precisely the number of solutions of the congruence $x^d \equiv 1 \pmod{p}$, which is d by theorem 5.78. Using a version of the Möbius inversion formula (see part 3 of remark 4.125), we obtain

$$f(d) = \sum_{e \mid d} \mu(e) \frac{d}{e} = \varphi(d),$$

as needed. The same argument is used in the next example.

Example 7.106. (Iran TST 2003) Let a_1, \dots, a_k be all primitive roots modulo an odd prime p . Prove that

$$a_1 + a_2 + \dots + a_k \equiv \mu(p-1) \pmod{p}.$$

Proof. For each $d \mid p-1$ set

$$f(d) = \sum_{x^d \equiv 1 \pmod{p}} x \pmod{p},$$

i.e. $f(d)$ is the remainder mod p of the sum of the solutions of the congruence $x^d \equiv 1 \pmod{p}$. By theorem 5.78 this congruence has d solutions, say x_1, \dots, x_d . Lagrange's theorem 5.69 yields

$$X^d - 1 \equiv (X - x_1)(X - x_2) \dots (X - x_d) \pmod{p},$$

and looking at the coefficient of X^{d-1} we obtain

$$f(d) \equiv x_1 + \dots + x_d \equiv 0 \pmod{p}$$

for $d > 1$. Thus $f(d) = 0$ for $d > 1$ and clearly $f(1) = 1$. On the other hand, it is clear that

$$f(d) = \sum_{u|d} g(u), \quad \text{where} \quad g(d) = \sum_{\text{ord}_p(x)=d} x \pmod{p}$$

is the remainder mod p of the sum of all numbers $x \in \{1, 2, \dots, p-1\}$ with $\text{ord}_p(x) = d$. Taking into account the values of f , the result follows by a version of the Möbius inversion formula (see remark 4.125). \square

The next example gives a different proof of the existence of primitive roots modulo p .

Example 7.107. a) Let n be a positive integer and let a_1, \dots, a_d be integers relatively prime to n . Prove that there is an integer c relatively prime to n such that

$$\text{ord}_n(c) = \text{lcm}(\text{ord}_n(a_1), \dots, \text{ord}_n(a_d)).$$

b) Deduce that there are primitive roots modulo p for any odd prime p .

Proof. a) Let $M = \text{lcm}(\text{ord}_n(a_1), \dots, \text{ord}_n(a_d))$ and assume that $M > 1$, the case $M = 1$ being clear. Let $M = p_1^{\alpha_1} \dots p_k^{\alpha_k}$ be the prime factorization of M and fix $i \in \{1, 2, \dots, k\}$. Since $p_i^{\alpha_i} \mid \text{lcm}(\text{ord}_n(a_1), \dots, \text{ord}_n(a_d))$, there is $x_i \in \{a_1, \dots, a_d\}$ such that $p_i^{\alpha_i} \mid \text{ord}_n(x_i)$. By proposition 7.66 the number

$c_i = x_i^{\frac{\text{ord}_n(x_i)}{p_i^{\alpha_i}}}$ has order $p_i^{\alpha_i}$ modulo n . Choosing $c = c_1 c_2 \dots c_k$ we obtain $\text{ord}_n(c) = M$. Indeed, it is clear that $c^M \equiv 1 \pmod{n}$, since $c_i^M \equiv 1 \pmod{n}$

for all i . On the other hand, if $c^N \equiv 1 \pmod{n}$, then $c^{\frac{N \cdot \frac{M}{p_i^{\alpha_i}}}{p_i^{\alpha_i}}} \equiv 1 \pmod{n}$,

which simplifies to $c_i^{\frac{N \cdot \frac{M}{p_i^{\alpha_i}}}{p_i^{\alpha_i}}} \equiv 1 \pmod{n}$. This yields $p_i^{\alpha_i} \mid N \cdot \frac{M}{p_i^{\alpha_i}}$, which in turn gives $p_i^{\alpha_i} \mid N$ for all i , thus $M \mid N$. We conclude that $\text{ord}_n(c) = M$, as needed.

b) Let

$$k = \text{lcm}(\text{ord}_p(1), \text{ord}_p(2), \dots, \text{ord}_p(p-1)).$$

By part a) we can write $k = \text{ord}_n(s)$ for some s relatively prime to p . By construction $a^k \equiv 1 \pmod{p}$ for all $a \in \{1, 2, \dots, p-1\}$, thus for all a relatively prime to p . Corollary 5.76 yields $p-1 \mid k$ and so $p-1 \mid \text{ord}_n(s) \mid p-1$. It follows that s is a primitive root mod p . \square

We end this section explaining the proof of theorem 7.103. The key technical ingredient is given by the following result, which is a simple consequence of the lifting the exponent lemma, more precisely of proposition 7.68.

Theorem 7.108. *Let p be an odd prime and let a be a primitive root mod p .*

a) a is a primitive root modulo p^2 if and only if $v_p(a^{p-1} - 1) = 1$.

b) If a is a primitive root mod p^2 , then a is a primitive root mod p^n for all $n \geq 1$.

Proof. a) Proposition 7.68 gives

$$\text{ord}_{p^2}(a) = (p-1) \cdot p^{2-v},$$

where $v = v_p(a^{p-1} - 1)$. Since a is a primitive root mod p^2 if and only if $\text{ord}_{p^2}(a) = p(p-1)$, the result follows.

b) This follows immediately from proposition 7.68 and part a). \square

Remark 7.109. Suppose that $a \in \{1, 2, \dots, p-1\}$ is a primitive root modulo p . It can (rarely) happen that $p^2 \mid a^{p-1} - 1$, in other words it is not true in general that a is a primitive root modulo p^2 . For instance one can prove that 5 is a primitive root modulo $p = 40487$ and $5^{p-1} \equiv 1 \pmod{p^2}$.

We can now easily finish the proof of theorem 7.103. We need to prove that there are primitive roots modulo p^n and $2p^n$ for any odd prime p and any $n \geq 1$. Choose a primitive root a modulo p and observe that $a+p$ is also a primitive root modulo p . We claim that one of the numbers a and $a+p$ is a primitive root modulo p^2 . Indeed, if neither of them is then the previous theorem yields $p^2 \mid a^{p-1} - 1$ and $p^2 \mid (a+p)^{p-1} - 1$. Using the binomial formula we obtain

$$(a+p)^{p-1} - 1 \equiv a^{p-1} - 1 + \binom{p-1}{1} a^{p-2} p \pmod{p^2}$$

and we conclude that $p^2 \mid (p-1)pa^{p-2}$, which is clearly impossible. This proves the existence of a primitive root b (equal to a or $a+p$) which is also a primitive root mod p^2 . Then b is a primitive root mod p^n for all $n \geq 1$ by the previous theorem. Finally, note that one of the numbers b and $b+p^n$ is odd, thus we may assume (possibly by replacing b with $b+p^n$) that b is odd. Since $\varphi(2p^n) = \varphi(p^n)$ and since $\text{ord}_{p^n}(b) = \varphi(p^n)$, we clearly have $\text{ord}_{2p^n}(b) = \varphi(2p^n)$ and so b is a primitive root mod $2p^n$. Theorem 7.103 is thus proved.

We end this rather long section with a few concrete examples in which the concept of primitive roots modulo n plays a key role.

Example 7.110. a) Prove that an odd prime p is congruent to 1 mod 8 if and only if the congruence $x^4 \equiv -1 \pmod{p}$ has solutions.

b) Deduce that if $p \equiv 1 \pmod{8}$ then $2^{\frac{p-1}{2}} \equiv 1 \pmod{p}$.

Proof. a) If $p \equiv 1 \pmod{8}$, take $x = g^{\frac{p-1}{8}}$ with g a primitive root mod p . Then $\text{ord}_p(x) = 8$, thus $x^8 \equiv 1 \pmod{p}$ and x^4 is not congruent to 1 mod p . It follows that $x^4 \equiv -1 \pmod{p}$, which proves one implication. Conversely, suppose that there is x such that $x^4 \equiv -1 \pmod{p}$. Then $\text{ord}_p(x) = 8$ since $\text{ord}_p(x)$ divides 8 and does not divide 4. Since $\text{ord}_p(x) \mid p-1$, we have $p \equiv 1 \pmod{8}$ and we are done.

b) Take x such that $x^4 \equiv -1 \pmod{p}$. Then $\gcd(p, x) = 1$, so there is an integer y such that $xy \equiv 1 \pmod{p}$. Let $z = x + y$, then $z^2 \equiv 2 + x^2 + y^2 \pmod{p}$. On the other hand $x^4 y^2 \equiv -y^2 \pmod{p}$ and $x^4 y^2 \equiv x^2 \pmod{p}$, thus $p \mid x^2 + y^2$ and so $z^2 \equiv 2 \pmod{p}$. It follows that $1 \equiv z^{p-1} \equiv 2^{\frac{p-1}{2}} \pmod{p}$ and we are done. \square

The next example gives a very conceptual proof of corollary 5.77.

Example 7.111. Prove that for all primes p and all positive integers n we have $1^n + 2^n + \dots + (p-1)^n \equiv 0 \pmod{p}$ if $p-1$ does not divide n , and $1^n + \dots + (p-1)^n \equiv -1 \pmod{p}$ if $p-1 \mid n$.

Proof. If a is a primitive root mod p , then $1, 2, \dots, p-1$ are congruent mod p to a permutation of $1, a, \dots, a^{p-2}$, hence

$$1^n + 2^n + \dots + (p-1)^n \equiv 1 + a^n + a^{2n} + \dots + a^{(p-2)n}.$$

The last expression can be easily computed: if $p-1 \mid n$, then $a^n \equiv 1 \pmod{p}$, hence $1^n + \dots + (p-1)^n \equiv p-1 \equiv -1 \pmod{p}$, while if $p-1$ does not divide n , then a^n is not congruent to 1 mod p , and

$$(a^n - 1)(1 + a^n + \dots + a^{(p-2)n}) = a^{(p-1)n} - 1 \equiv 0 \pmod{p},$$

hence $1 + a^n + \dots + a^{(p-2)n} \equiv 0 \pmod{p}$ and we are done. \square

Example 7.112. Let a, n, k be integers with $n, k > 0$ and $\gcd(a, n) = 1$. Suppose that there are primitive roots mod n , and let $d = \gcd(k, \varphi(n))$.

a) Prove that the congruence $x^k \equiv a \pmod{n}$ has solutions if and only if $a^{\frac{\varphi(n)}{d}} \equiv 1 \pmod{n}$, and in this case the congruence has d solutions.

b) For how many integers $a \in \{0, 1, \dots, n-1\}$ relatively prime to n does the congruence $x^k \equiv a \pmod{n}$ have solutions?

Proof. a) Let g be a primitive root modulo n . If $x^k \equiv a \pmod{n}$, then $\gcd(x, n) = 1$ since $\gcd(a, n) = 1$. Thus we can write $x \equiv g^j \pmod{n}$ and $a \equiv g^u \pmod{n}$ for unique integers $j, u \in \{0, 1, \dots, \varphi(n) - 1\}$. The congruence $x^k \equiv a \pmod{n}$ is then equivalent to $g^{kj-u} \equiv 1 \pmod{n}$ and to $kj \equiv u \pmod{\varphi(n)}$. This linear congruence (the "unknown" being j) has solutions if and only if u is a multiple of d , and if this is the case the congruence has exactly d solutions. On the other hand, the congruence $a^{\frac{\varphi(n)}{d}} \equiv 1 \pmod{n}$ is equivalent to $g^{u\frac{\varphi(n)}{d}} \equiv 1 \pmod{n}$, then to $u\frac{\varphi(n)}{d} \equiv 0 \pmod{\varphi(n)}$ and finally to $u \equiv 0 \pmod{d}$. The result follows.

b) By a) we need to find the number of solutions of the congruence $a^{\frac{\varphi(n)}{d}} \equiv 1 \pmod{n}$. By the proof of part a), this amounts to finding the number of integers $u \in \{0, 1, \dots, \varphi(n) - 1\}$ such that $u \equiv 0 \pmod{d}$, which is exactly $\frac{\varphi(n)}{d}$. \square

Remark 7.113. By taking $n = p$ an odd prime and $k = 2$, we recover Euler's criterion and the formula for the number of quadratic residues mod p .

Example 7.114. Prove that the number of solutions of the congruence $x^{n-1} \equiv 1 \pmod{n}$ is $\prod_{p \mid n} \gcd(p-1, n-1)$.

Proof. Let $n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$ be the prime factorization of n and let a_i be the number of solutions of the congruence $x^{n-1} \equiv 1 \pmod{p_i^{\alpha_i}}$. By the Chinese remainder theorem (more precisely by theorem 7.9), it suffices to prove that $\prod_{i=1}^k a_i = \prod_{i=1}^k \gcd(p_i - 1, n - 1)$. We will prove that $a_i = \gcd(p_i - 1, n - 1)$ for $1 \leq i \leq k$. If $p_i > 2$ then $p_i^{\alpha_i}$ has primitive roots and so (by the previous example) the congruence $x^{n-1} \equiv 1 \pmod{p_i^{\alpha_i}}$ has

$$\gcd(n - 1, \varphi(p_i^{\alpha_i})) = \gcd(n - 1, p_i^{\alpha_i-1}(p_i - 1)) = \gcd(n - 1, p_i - 1)$$

solutions, as desired. A similar argument works when $p_i = 2$ and $\alpha_i > 2$. Suppose that $p_i = 2$ and $\alpha_i \in \{1, 2\}$, we need to prove that the congruences $x^{n-1} \equiv 1 \pmod{2}$ and $x^{n-1} \equiv 1 \pmod{4}$ have exactly one solution when $v_2(n) = 1$ and $v_2(n) = 2$ respectively. This is clear. \square

Example 7.115. (AMM E 3212) Is it true that if n is sufficiently large and a_1, a_2, \dots, a_n is an arbitrary permutation of $1, 2, \dots, n$, then we can find integers i, d such that $1 \leq i < i + d < i + 2d \leq n$ and a_i, a_{i+d}, a_{i+2d} form an arithmetic progression?

Proof. The answer is negative. If p is an odd prime, let g be a primitive root mod p and consider the permutation a_1, \dots, a_{p-1} of $1, 2, \dots, p - 1$ defined by $a_i \equiv g^i \pmod{p}$. If a_i, a_{i+d}, a_{i+2d} form an arithmetic progression, then $g^i + g^{i+2d} \equiv 2g^{i+d} \pmod{p}$ and so $(g^d - 1)^2 \equiv 0 \pmod{p}$. This forces $g^d \equiv 1 \pmod{p}$, hence $p - 1 \mid d$ and $d \geq p - 1$, a contradiction. \square

Example 7.116. (Kömal) Is there a positive integer n such that every nonzero digit (in base ten) appears the same number of times in the decimal representation of each of the numbers $n, 2n, 3n, \dots, 2016n$?

Proof. Suppose that there is a prime $p > 2016$ such that 10 is a primitive root modulo p . Consider an integer n such that $n \cdot p = 10^{p-1} - 1$. Arguing as in the proof of example 7.75, we see that the periods of the fractions $\frac{1}{p}, \frac{2}{p}, \dots, \frac{2016}{p}$ are obtained by cyclic permutations of the period of $\frac{1}{p}$, and the decimal representations of the numbers $n, 2n, 3n, \dots, 2016n$ are also obtained by cyclic permutations of the digits (with an appropriate number of leading zeroes), hence n is a solution of the problem.

We prove now the existence of such a prime p . We will check that $p = 2^{16} + 1$ works. It is well-known (and not difficult to prove) that p is indeed a prime. The order of 10 modulo p divides $p - 1 = 2^{16}$, thus if the order is not $p - 1$, then it must divide $\frac{p-1}{2}$ and so $10^{\frac{p-1}{2}} \equiv 1 \pmod{p}$. It follows that $\left(\frac{2}{p}\right) \cdot \left(\frac{5}{p}\right) = \left(\frac{10}{p}\right) = 1$, which is impossible, since $\left(\frac{2}{p}\right) = 1$ (as $p \equiv 1 \pmod{8}$) and $\left(\frac{5}{p}\right) = \left(\frac{p}{5}\right) = \left(\frac{2}{5}\right) = -1$ (we used here the quadratic reciprocity law and the fact that $p \equiv 2 \pmod{5}$). \square

Remark 7.117. It is not known whether 10 is a primitive root modulo p for infinitely many primes p . This is a special case of a famous conjecture due to Artin, stating that any integer $a \neq -1$ which is not a perfect square is a primitive root modulo p for infinitely many primes p .

Example 7.118. (USA TST 2010) Is there a positive integer k such that $p = 6k + 1$ is a prime and $\binom{3k}{k} \equiv 1 \pmod{p}$?

Proof. The answer is negative. Suppose that $p = 6k + 1$ is a prime and $\binom{3k}{k} \equiv 1 \pmod{p}$. Let g be a primitive root mod p and let $z = g^6$. Then z has order k mod p , hence $\sum_{i=0}^{k-1} z^{ij}$ is 0 modulo p , unless j is a multiple of k . We deduce that

$$\begin{aligned} S &= \sum_{i=0}^{k-1} (1 + z^i)^{3k} = \sum_{i=0}^{k-1} \left(\sum_{j=0}^{3k} \binom{3k}{j} z^{ij} \right) = \sum_{j=0}^{3k} \binom{3k}{j} \sum_{i=0}^{k-1} z^{ij} \\ &\equiv \left(\binom{3k}{0} + \binom{3k}{k} + \binom{3k}{2k} + \binom{3k}{3k} \right) k \\ &= \left(2 + 2 \binom{3k}{k} \right) k \equiv 4k \pmod{p}. \end{aligned}$$

On the other hand, for all $0 \leq i \leq k - 1$ we have

$$(1 + z^i)^{3k} \equiv (1 + z^i)^{\frac{p-1}{2}} \equiv -1, 0, 1 \pmod{p}.$$

However we cannot have k remainders mod p , each of them $-1, 0$ or 1 , adding up to $4k$ modulo p . The result follows. \square

7.4 Problems for practice

The Chinese remainder theorem

1. (Poland 2003) A polynomial f with integer coefficients has the property that $\gcd(f(a), f(b)) = 1$ for some integers $a \neq b$. Prove that there is an infinite set of integers S such that $\gcd(f(m), f(n)) = 1$ whenever m, n are distinct elements of S .
2. Prove that for all positive integers k and n there exists a set S of n consecutive positive integers such that each $x \in S$ has at least k distinct prime divisors that do not divide any other element of S .
3. A lattice point is called visible if its coordinates are relatively prime integers. Prove that for any positive integer k there is a lattice point whose distance from each visible lattice point is greater than k .
4. a) Prove that for all $n \geq 1$ there is a positive integer a such that $a, 2a, \dots, na$ are all perfect powers.
b) (Balkan 2000) Prove that for all $n \geq 1$ there is a set A of n positive integers such that for all $1 \leq k \leq n$ and all $x_1, x_2, \dots, x_k \in A$ the number $\frac{x_1 + x_2 + \dots + x_k}{k}$ is a perfect power.
5. Let a, b, c be pairwise distinct positive integers. Prove that there is an integer n such that $a + n, b + n, c + n$ are pairwise relatively prime.
6. (AMM) Prove that there are arbitrarily long sequences of consecutive integers, none of which can be written as the sum of two perfect squares.
7. Let f be a nonconstant polynomial with integer coefficients and let n and k be positive integers. Prove that there is a positive integer a such that each of the numbers $f(a), f(a+1), \dots, f(a+n-1)$ has at least k distinct prime divisors.
8. (IMC 2013) Let p and q be relatively prime positive integers. Prove that

$$\sum_{k=0}^{pq-1} (-1)^{\lfloor \frac{k}{p} \rfloor + \lfloor \frac{k}{q} \rfloor} = \begin{cases} 0 & \text{if } pq \text{ is even} \\ 1 & \text{if } pq \text{ odd} \end{cases}$$

9. (IMO 1999 Shortlist) Find all positive integers n for which there is an integer m such that $2^n - 1 \mid m^2 + 9$.
10. (Bulgaria 2003) A finite set C of positive integers is called good if for any $k \in \mathbf{Z}$ there exist $a \neq b \in C$ such that $\gcd(a+k, b+k) > 1$. Prove that if the sum of the elements of a good set C equals 2003, then there exists $c \in C$ such that the set $C - \{c\}$ is good.
11. Is there a sequence of 101 consecutive odd integers such that each term of the sequence has a prime factor not exceeding 103?
12. (USA TST 2010) The sequence $(a_n)_{n \geq 1}$ satisfies $a_1 = 1$ and

$$a_n = a_{\lfloor n/2 \rfloor} + a_{\lfloor n/3 \rfloor} + \dots + a_{\lfloor n/n \rfloor} + 1$$

for all $n \geq 2$. Prove that $a_n \equiv n \pmod{2^{2010}}$ for infinitely many n .

13. (China TST 2014) A function $f: \mathbf{N} \rightarrow \mathbf{N}$ satisfies for all $m, n \geq 1$

$$\gcd(f(m), f(n)) \leq \gcd(m, n)^{2014} \quad \text{and} \quad n \leq f(n) \leq n + 2014.$$

Prove that there is a positive integer N such that $f(n) = n$ for $n \geq N$.

Euler's theorem

14. (Iran 2007) Let n be a positive integer such that $\gcd(n, 2(2^{1386} - 1)) = 1$. Let $a_1, a_2, \dots, a_{\varphi(n)}$ be a reduced residue system modulo n . Prove that

$$n \mid a_1^{1386} + a_2^{1386} + \dots + a_{\varphi(n)}^{1386}$$

15. Let $n > 1$ be an integer and let $r_1, r_2, \dots, r_{\varphi(n)}$ be a reduced residue system modulo n . For which integers a is $r_1 + a, r_2 + a, \dots, r_{\varphi(n)} + a$ a reduced residue system modulo n ?
16. Prove that any positive integer n has a multiple whose sum of digits is n .
17. For which integers $n > 1$ is there a polynomial f with integer coefficients such that $f(k) \equiv 0 \pmod{n}$ or $f(k) \equiv 1 \pmod{n}$ for any integer k , and both these congruences have solutions?

18. (Saint Petersburg 1998) Is there a nonconstant polynomial f with integer coefficients and an integer $a > 1$ such that the numbers $f(a)$, $f(a^2)$, $f(a^3)$, ... are pairwise relatively prime?
19. a) (IMO 1971) Prove that the sequence $(2^n - 3)_{n \geq 1}$ contains an infinite subsequence in which every two distinct terms are relatively prime.
b) (Romania TST 1997) Let $a > 1$ be a positive integer. Prove the same result as in a) for the sequence $(a^{n+1} + a^n - 1)_{n \geq 1}$.
20. (China TST 2005) Integers a_0, a_1, \dots, a_n and x_0, x_1, \dots, x_n satisfy

$$a_0 x_0^k + a_1 x_1^k + \dots + a_n x_n^k = 0$$

for all $1 \leq k \leq r$, where r is a positive integer. Prove that m divides $a_0 x_0^m + a_1 x_1^m + \dots + a_n x_n^m$ for all $r + 1 \leq m \leq 2r + 1$.

21. (Hong Kong 2010) Let n be an integer greater than 1 and let $1 \leq a_1 < \dots < a_k \leq n$ be the totatives of n . Prove that for any integer a relatively prime to n we have

$$\frac{a^{\phi(n)} - 1}{n} \equiv \sum_{i=1}^k \frac{1}{aa_i} \left[\frac{aa_i}{n} \right] \pmod{n}$$

22. (Kömal) Let x_1, x_2, \dots, x_n be integers such that $\gcd(x_1, \dots, x_n) = 1$. Prove that if $s_i = x_1^i + x_2^i + \dots + x_n^i$, then

$$\gcd(s_1, s_2, \dots, s_n) \mid \text{lcm}(1, 2, \dots, n).$$

23. (Brazil 2005) Let a and c be positive integers. Prove that for any integer b there is a positive integer x such that

$$a^x + x \equiv b \pmod{c}.$$

24. (Ibero American 2012) Prove that for any integer $n > 1$ there exist n consecutive positive integers such that none of them is divisible by the sum of its digits.

Order modulo n

25. (Russia 2006) Let x and y be purely periodic decimal fractions such that $x + y$ and xy are purely periodic decimal fractions with period length T . Prove that the lengths of the periods of x and y are not greater than T .
26. (Iran 2013) Let p be an odd prime and let d be a positive divisor of $p - 1$. Let S be the set of integers $x \in \{1, 2, \dots, p - 1\}$ for which the order of x modulo p is d . Find the remainder of $\prod_{x \in S} x$ when divided by p .
27. Let a, b, n be positive integers with $a \neq b$. Prove that

$$2n \mid \varphi(a^n + b^n) \quad \text{and} \quad n \mid \varphi\left(\frac{a^n - b^n}{a - b}\right).$$

28. Find all primes p and q such that $p^2 + 1 \mid 2003^q + 1$ and $q^2 + 1 \mid 2003^p + 1$.
29. (MOSP 2001) Let p be a prime number and let m, n be integers greater than 1 such that $n \mid m^{p(n-1)} - 1$. Prove that $\gcd(m^{n-1} - 1, n) > 1$.
30. a) (Pepin's test) Let n be a positive integer and let $k = 2^{2^n} + 1$. Prove that k is a prime if and only if $k \mid 3^{\frac{k-1}{2}} + 1$.
- b) (Euler-Lagrange) Let $p \equiv -1 \pmod{4}$ be a prime. Prove that $2p + 1$ is a prime if and only if $2p + 1 \mid 2^p - 1$.
31. Let $p > 2$ be an odd prime and let a be a primitive root modulo p . Prove that $a^{\frac{p-1}{2}} \equiv -1 \pmod{p}$.
32. Suppose that $n > 1$ is an integer for which there are primitive roots modulo n . Prove that the set $\{1, 2, \dots, n\}$ contains exactly $\varphi(\varphi(n))$ primitive roots modulo n .
33. Let p be an odd prime. Prove that p is a Fermat prime (i.e. of the form $2^n + 1$ with $n \geq 1$) if and only if every quadratic non-residue mod p is a primitive root mod p .

34. Let $\lambda(n)$ be the least positive integer k such that $x^k \equiv 1 \pmod{n}$ for all x relatively prime to n . Prove that
- a) If k is a positive integer such that $x^k \equiv 1 \pmod{n}$ for all x relatively prime to n , then k is a multiple of $\lambda(n)$.
 - b) $\lambda(mn) = \text{lcm}(\lambda(m), \lambda(n))$ for m, n relatively prime.
 - c) We have $\lambda(n) = \varphi(n)$ when $n = 2, 4$ or a power of an odd prime, and $\lambda(2^n) = 2^{n-2}$ for $n \geq 3$.
 - d) For each n , the set of numbers $\text{ord}_n(x)$ (over all x relatively prime to n) is precisely the set of positive divisors of $\lambda(n)$.
35. Let $p > 2$ be a prime and let a be a primitive root mod p . Prove that $-a$ is a primitive root mod p if and only if $p \equiv 1 \pmod{4}$.
36. (Unesco Competition 1995) Let m, n be integers greater than 1. Prove that the remainders of the numbers $1^n, 2^n, \dots, m^n$ modulo m are pairwise distinct if and only if m is square-free and n is relatively prime to $\varphi(m)$.
37. (adapted from Tuymaada 2011) Prove that among 2500 consecutive positive integers there is an integer n such that the length of the period of the decimal expansion of $\frac{1}{n}$ is greater than 2011.
38. Is there a positive integer which is divisible by the product P of its digits and such that P is a power of 7 greater than 10^{2016} ?
39. Let m, n be positive integers. Prove that there is a positive integer k such that $2^k \equiv 1999 \pmod{3^m}$ and $2^k \equiv 2009 \pmod{5^n}$.
40. (Iran 2012) Let p be an odd prime. Prove that there is a positive integer x such that x and $4x$ are both primitive roots modulo p .
41. (Brazil 2009) Let p, q be odd primes such that $q = 2p + 1$. Prove that there is a multiple of q whose sum of digits is 1, 2 or 3.
42. (Brazil 2012) Find the least positive integer n for which there is a positive integer k such that the last 2012 decimal digits of n^k are all 1's.

43. (Nieuw Archief voor Wiskunde) Suppose that $\alpha \geq \frac{\log 10}{\log 5} = 1.43067\dots$. Prove that for any $n \geq 1$, any sequence of n digits (between 0 and 9) occurs as a sequence of consecutive digits in the last $\lceil \alpha n \rceil$ digits of some power of 2.
44. Find all sequences of positive integers $(a_n)_{n \geq 1}$ such that
- a) $m - n \mid a_m - a_n$ for all positive integers m, n ;
 - b) If m, n are relatively prime, then a_m and a_n are relatively prime.
45. (adapted after China TST 2012) Let $n > 1$ be an integer. Find all functions $f : \mathbf{Z} \rightarrow \{1, 2, \dots, n\}$ such that for each $k \in \{1, 2, \dots, n-1\}$ there is $j(k) \in \mathbf{Z}$ such that for all integers m we have

$$f(m + j(k)) \equiv f(m + k) - f(m) \pmod{n+1}.$$

Chapter 8

Solutions to practice problems

8.1 Divisibility

1. Prove that the last $n + 2$ digits of 5^{2^n+n+2} are the digits of 5^{n+2} , completed on the left with some zeros.

Proof. This is equivalent to the congruence

$$5^{2^n+n+2} \equiv 5^{n+2} \pmod{10^{n+2}}.$$

Thus it suffices to show that $5^{2^n} \equiv 1 \pmod{2^{n+2}}$. This follows from theorem 2.31, i.e. from the equality

$$5^{2^n} - 1 = 2^2 \cdot (5 + 1)(5^2 + 1) \dots (5^{2^{n-1}} + 1). \quad \square$$

2. Is there a polynomial f with integer coefficients such that the congruence $f(x) \equiv 0 \pmod{6}$ has 2, 3 as solutions, but no other solution in the set $\{0, 1, \dots, 5\}$?

Proof. The answer is negative. Indeed, suppose that f is such a polynomial, then $3f(2) - 2f(3)$ is a multiple of 6.

On the other hand $f(2) \equiv f(0) \pmod{2}$, thus $3f(2) \equiv 3f(0) \pmod{6}$. Similarly $2f(3) \equiv 2f(0) \pmod{6}$, thus $3f(2) - 2f(3) \equiv f(0) \pmod{6}$ and so $6 \mid f(0)$, a contradiction. \square

3. (Iran 2003) Is there an infinite set S such that for all distinct elements a, b of S we have $a^2 - ab + b^2 \mid a^2b^2$?

Proof. There is no such set S . Assuming that S exists, fix $a \in S$ and choose any $b > a$ in S . Then $a^2 - ab + b^2 \mid a^2b^2$, but $a^2 - ab + b^2 \nmid a^2(a^2 - ab + b^2)$. Taking the difference, we deduce that

$$a^2 - ab + b^2 \mid a^3b - a^4$$

and so

$$a^2 - ab + b^2 \leq a^3b - a^4 < a^3b.$$

Since the left-hand side is greater than or equal to $\frac{3b^2}{4}$, we conclude that $b < \frac{4a^3}{3}$. Since $b > a$ was arbitrary in S , we conclude that S is finite, a contradiction. \square

4. (Russia 2003) Is it possible to write a positive integer in every cell of an infinite chessboard in such a manner that for all integers $m, n > 100$, the sum of numbers in every $m \times n$ rectangle is divisible by $m + n$?

Proof. The answer is negative: assume that we managed to write positive integers as in the statement of the problem and choose any integer $n > 100$, as well as an arbitrary cell of the chessboard.

Consider the $(2n + 1) \times (2n + 1)$ square centered at that cell. One can partition this square into four $n \times (n + 1)$ or $(n + 1) \times n$ rectangles R_1, \dots, R_4 , plus the central cell. By hypothesis the sum of the entries in the cells of R_i is a multiple of $2n + 1$ for $1 \leq i \leq 4$. Also, the sum of the entries in the cells of the square is a multiple of $4n + 2$, thus a multiple of $2n + 1$. It follows that the number in the central cell is a multiple of $2n + 1$. Thus the number in each cell is a multiple of $2n + 1$, and this for all $n > 100$. It follows that all numbers in the cells are 0, a contradiction. \square

5. Prove that if $k > 1$ is an integer then there are infinitely many positive integers n such that $n|k^n + 1$.

Proof. If k is odd, then $n = 2$ is a solution, while if k is even, then $n = k + 1$ is a solution of the problem. Starting with a solution n we will create another one which is larger. Namely, let $n_1 = k^n + 1$, which is certainly larger than n . Let us check whether n_1 is a solution, i.e. whether $n_1 | k^{n_1} + 1$, or equivalently $k^n + 1 | k^{n_1} + 1$. This will happen if $\frac{n_1}{n}$ (which is an integer) is odd. This is automatic if k is even, as then n_1 is odd.

Things are a little bit more complicated when k is odd, as then $n_1 = k^n + 1$ is even, so it is not a priori clear that $\frac{n_1}{n}$ is odd. However, if we know that n is even, then $k^n + 1$ is not a multiple of 4 (as this is the case with any number of the form $x^2 + 1$), thus $\frac{n_1}{n}$ is odd and we are done.

The strategy is now clear: let $n_0 = k + 1$ when k is even, and $n_0 = 2$ when k is odd. Then, for $j \geq 0$, define $n_{j+1} = k^{n_j} + 1$. The previous discussion shows that n_0, n_1, \dots are all solutions of the problem. \square

6. (Kvant M 904) For each positive integer A with decimal representation

$$A = \overline{a_n a_{n-1} \dots a_0}$$

we set

$$F(A) = a_n + 2a_{n-1} + \dots + 2^{n-1}a_1 + 2^n a_0$$

and consider the sequence $A_0 = A$, $A_1 = F(A_0)$, $A_2 = F(A_1), \dots$

(i) Prove that there is a term A^* of this sequence such that $A^* < 20$ and $F(A^*) = A^*$.

(ii) Find A^* for $A = 19^{2013}$.

Proof. (i) If A is an one-digit number or $A = 19$, then one easily checks that $F(A) = A$. We will show that for any other A , $F(A) < A$. From this it follows that the sequence A_0, A_1, \dots is strictly decreasing until the number 19 or an one digit number appears. If we denote this number by A^* we have $F(A^*) = A^*$.

Suppose A has two digits and satisfies $F(A) \geq A$. Writing $A = 10a + b$, this becomes the inequality $10a + b \leq 2b + a$ or equivalently $9a \leq b$. Since a is nonzero, and b is a single digit, we have $9 \leq 9a \leq b \leq 9$, hence we must have equality throughout, thus $a = 1$ and $b = 9$ and $A = 19$. If A has $n + 1$ digits for some $n \geq 2$, then $A \geq 10^n$. Hence

$$\begin{aligned} F(A) &= a_n + 2a_{n-1} + \cdots + 2^{n-1}a_1 + 2^na_0 \leq 9 + 2 \cdot 9 + \cdots + 2^n \cdot 9 \\ &= 9(2^{n+1} - 1) < 72 \cdot 2^{n-2} < 10^n \leq A. \end{aligned}$$

Thus we have shown that $F(A) < A$ unless A is a single digit or $A = 19$, as desired.

(ii) Note that

$$2^n A - F(A) = (20^n - 1)a_n + 2(20^{n-1} - 1)a_{n-1} + \cdots + 2^{n-1}(20 - 1)a_1$$

is divisible by 19. So, if A is divisible by 19 then $F(A)$ is also divisible by 19 and therefore all terms of the sequence are divisible by 19. Now if $A = 19^{2013}$ then all terms of the sequence are divisible by 19 and therefore $A^* = 19$. \square

7. Are there infinitely many 5-tuples (a, b, c, d, e) of positive integers such that $1 < a < b < c < d < e$ and $a \mid b^2 - 1$, $b \mid c^2 - 1$, $c \mid d^2 - 1$, $d \mid e^2 - 1$ and $e \mid a^2 - 1$?

Proof. The answer is positive. The easiest way to ensure that $b \mid c^2 - 1$, $c \mid d^2 - 1$ and $d \mid e^2 - 1$ is to take $b = c - 1$, $c = d - 1$ and $d = e - 1$. This reduces the problem to finding infinitely many pairs (a, b) with $1 < a < b$ and $b + 3 \mid a^2 - 1$, $a \mid b^2 - 1$. Simply take $b = 2a - 1$ with $a > 1$ and observe that $a \mid b^2 - 1$ and $b + 3 = 2(a + 1) \mid a^2 - 1$ if a is odd. \square

8. (Romania JBMO TST 2003) Let A be a finite set of positive integers with at least three elements. Prove that there are two elements of A whose sum does not divide the sum of the other elements of A .

Proof. Let $a_1 < a_2 < \dots < a_k$ be the elements of A , and assume that $a_i + a_j$ divides $\sum_{l \neq i, j} a_l$ for all $i \neq j$. Then $a_i + a_j$ also divides $S = a_1 + a_2 + \dots + a_k$ for all $i \neq j$. In particular, there are positive integers x_i such that $S = x_i(a_k + a_i)$ for $1 \leq i < k$. Since $a_1 < a_2 < \dots < a_{k-1}$, it follows that $x_1 > x_2 > \dots > x_{k-1}$. Moreover $x_i > 1$ for all i , and $a_k x_i < S < k a_k$, thus $x_i < k$ for all i . It follows that $\{2, 3, \dots, k-1\}$ contains at least $k-1$ distinct positive integers x_1, x_2, \dots, x_{k-1} , a contradiction. \square

9. (Iran 2005) Prove that there are infinitely many positive integers n such that

$$n \mid 3^{n+1} - 2^{n+1}.$$

Proof. We will look for n of the form $3^a - 2^a$ for some $a > 1$. The condition $n \mid 3^{n+1} - 2^{n+1}$ is satisfied if $a \mid n+1 = 3^a - 2^a + 1$. We claim that $a = 2 \cdot 3^k$ works for all $k \geq 1$. It suffices to prove that $3^k \mid 4^{3^k} - 1$. But

$$4^{3^k} - 1 = (4 - 1)(4^2 + 4 + 1)(4^{2 \cdot 3} + 4^3 + 1) \dots (4^{2 \cdot 3^{k-1}} + 4^{3^{k-1}} + 1)$$

and each of the factors in the above product is a multiple of 3. \square

10. (Mathematical Reflections S 259) Let a, b, c, d, e be integers such that

$$a(b+c) + b(c+d) + c(d+e) + d(e+a) + e(a+b) = 0.$$

Prove that $a+b+c+d+e$ divides $a^5 + b^5 + c^5 + d^5 + e^5 - 5abcde$.

Proof. Let A, B, C, D, E be integers such that

$$(X-a)(X-b)(X-c)(X-d)(X-e) = X^5 + AX^4 + BX^3 + CX^2 + DX + E$$

as polynomials. Thus

$$A = -(a+b+c+d+e), \quad B = ab+ac+ad+ae+\dots+de, \quad \dots, \quad E = -abcde.$$

Note that

$$B = a(b+c) + b(c+d) + c(d+e) + d(e+a) + e(a+b) = 0$$

by hypothesis. For each $x \in \{a, b, c, d, e\}$ we have

$$x^5 + Ax^4 + Cx^2 + Dx + E = 0.$$

Adding these 5 equations yields

$$\begin{aligned} & a^5 + b^5 + c^5 + d^5 + e^5 - 5abcde \\ & + A(a^4 + \dots + e^4) + C(a^2 + \dots + e^2) + D(a + \dots + e) = 0. \end{aligned}$$

Since the last term of the sum is a multiple of A , as is $A(a^4 + \dots + e^4)$, it suffices to prove that $C(a^2 + b^2 + c^2 + d^2 + e^2)$ is a multiple of A . But

$$A^2 = (a+b+c+d+e)^2 = a^2 + b^2 + c^2 + d^2 + e^2 + 2B = a^2 + b^2 + c^2 + d^2 + e^2,$$

yielding the desired result. \square

11. (Kazakhstan 2011) Find the smallest integer $n > 1$ such that there exist positive integers a_1, a_2, \dots, a_n for which

$$a_1^2 + \dots + a_n^2 \mid (a_1 + \dots + a_n)^2 - 1.$$

Proof. Let n be a solution of the problem and write

$$(a_1 + \dots + a_n)^2 - 1 = k(a_1^2 + \dots + a_n^2) \quad (1)$$

for some positive integer k . We claim that $a_1 + \dots + a_n$ is odd. Assume for contradiction that this is not the case. Since

$$a_1^2 + \dots + a_n^2 - (a_1 + \dots + a_n) = a_1(a_1 - 1) + \dots + a_n(a_n - 1) \quad (2)$$

is even and since $a_1 + \dots + a_n$ is even, we deduce that $a_1^2 + \dots + a_n^2$ is even, which contradicts relation (1) (the left-hand side is odd, while the right-hand side is even). Thus $a_1 + \dots + a_n$ is odd and so $(a_1 + \dots + a_n)^2 - 1$ is a multiple of 8. Relation (2) combined with the fact that $a_1 + \dots + a_n$ is odd shows that $a_1^2 + \dots + a_n^2$ is odd too. Since $k(a_1^2 + \dots + a_n^2)$ is a multiple of 8 and $a_1^2 + \dots + a_n^2$ is odd, we deduce that k is a multiple of 8 and so $k \geq 8$. On the other hand, the Cauchy-Schwarz inequality yields

$$k(a_1^2 + \dots + a_n^2) = (a_1 + \dots + a_n)^2 - 1 \leq n(a_1^2 + \dots + a_n^2) - 1 < n(a_1^2 + \dots + a_n^2).$$

We deduce that $n > 8$ and so the smallest solution of the problem is at least 9. To see that this is indeed the solution, choose $a_1 = a_2 = 2$, $a_3 = \dots = a_9 = 1$. \square

12. (Kvant 898) Find all odd integers $0 < a < b < c < d$ such that

$$ad = bc, \quad a + d = 2^k, \quad b + c = 2^m$$

for some positive integers k and m .

Proof. We first prove that $k > m$. Indeed, we have

$$2^k - 2^m = a - b + d - \frac{ad}{b} = \frac{(b-a)(d-b)}{b} > 0.$$

Next we prove that $a + b = 2^{m-1}$. To do this we write the identity $ad = bc$ as $a(2^k - a) = b(2^m - b)$, i.e. $b^2 - a^2 = 2^m(b - 2^{k-m}a)$. Hence 2^m divides $(b-a)(b+a)$ and since b and a are odd one of $b-a$ and $b+a$ is divisible by 2 and the other by 2^{m-1} . But

$$b - a < b < \frac{b+c}{2} = 2^{m-1}$$

and therefore $b + a$ is divisible by 2^{m-1} .

On the other hand $b+a < b+c = 2^m$ and we conclude that $b+a = 2^{m-1}$. Hence $b = 2^{m-1} - a$, $c = 2^m - b = 2^{m-1} + a$ and $ad = bc = 2^{2m-2} - a^2$. This shows that a divides 2^{2m-2} and therefore $a = 1$ since a is odd. Thus $a = 1, b = 2^{m-1} - 1, c = 2^{m-1} + 1, d = 2^{2m-2} - 1, k = 2m - 2$, where $m \geq 3$ is an arbitrary integer. \square

13. f is a polynomial with integer coefficients such that $f(n) > n$ for every positive integer n . Define a sequence $(x_n)_{n \geq 1}$ by $x_1 = 1$ and $x_{i+1} = f(x_i)$. Assuming that each positive integer has a multiple among x_1, x_2, \dots , prove that $f(X) = X + 1$.

Proof. By hypothesis we have $x_{i+1} > x_i$ for all $i \geq 1$, that is, the sequence $(x_n)_{n \geq 1}$ is increasing. Moreover, again by hypothesis given $n \geq 2$ we can find m such that $x_n - x_{n-1} \mid x_m$. Choose a minimal such m and suppose that $m \geq n$. Let us note that $x_{j+2} - x_{j+1} = f(x_{j+1}) - f(x_j)$ is a multiple of $x_{j+1} - x_j$ for all j , therefore $x_{j+1} - x_j \mid x_{k+1} - x_k$ if $k \geq j$. Thus $x_n - x_{n-1} \mid x_m - x_{m-1}$ and so $x_n - x_{n-1} \mid x_{m-1}$, contradicting the minimality of m . Thus $m < n$ and so $x_n - x_{n-1} \leq x_{m-1}$. We conclude that $f(x_{n-1}) \leq 2x_{n-1}$ for all $n \geq 2$. If $\deg f \geq 2$, then for x large enough we have $f(x) > 2x$, which contradicts the previous inequality. Thus $f(X) = aX + b$ for some integers a, b . Since $f(n) > n$ for all n , we deduce that $a \geq 1$. Since $ax_{n-1} + b \leq 2x_{n-1}$ for all $n > 1$, we have $a \leq 2$. Thus $a = 1$ or $a = 2$. If $a = 1$, then $x_n = 1 + (n-1)b$ and by assumption there is n such that $b \mid x_n$, which yields $b = 1$ and $f(X) = X + 1$. If $a = 2$, an easy induction shows that $x_n = 2^{n-1}(1+b) - b$. By assumption there is n such that $1+b \mid x_n$, which forces $1+b \mid b$ and then $1+b \mid 1$. This can only happen if $1+b = 1$, i.e. $b = 0$ and hence $x_n = 2^{n-1}$. But then trivially 3 does not divide any term of the sequence, a contradiction. \square

14. (Iran 2013) Suppose that a, b are two odd positive integers such that $2ab + 1 \mid a^2 + b^2 + 1$. Prove that $a = b$.

Proof. Arguing as usual by infinite descent, we consider a pair (a, b) satisfying the hypothesis of the problem and failing to satisfy the conclusion, such that $a + b$ has the smallest possible value. We may assume that $a > b$. Write

$$a^2 + b^2 + 1 = c(2ab + 1)$$

and note that $c \neq 1$, since $a \neq b$. Consider the other solution

$$a' = 2bc - a = \frac{b^2 + 1 - c}{a}$$

of the equation

$$x^2 - 2bcx + b^2 + 1 - c = 0.$$

Note that $a' = 2bc - a$ is odd and $a' \neq b$ (since $c \neq 1$ and $(a')^2 + b^2 + 1 = c(2a'b + 1)$). Also, note that $a' > 0$, since otherwise $a' \leq -1$ and so

$b^2 + 1 - c \leq -a$, thus

$$b^2 + a + 1 \leq c = \frac{a^2 + b^2 + 1}{2ab + 1} < \frac{2a^2 + b^2}{2ab} \leq a + b^2,$$

a contradiction. By minimality of (a, b) , we obtain $a' \geq a$. This is however impossible, since (recall that $c > 1$)

$$a' = \frac{b^2 + 1 - c}{a} < \frac{b^2}{a} \leq a.$$

The result follows. \square

15. (Kvant) Prove that $n^2 + 1$ divides $n!$ for infinitely many positive integers n .

Proof. We start by choosing n so that $n^2 + 1$ admits a nontrivial factorization. For instance, choosing $n = 2k^2$ yields

$$n^2 + 1 = 4k^4 + 1 = (2k^2)^2 + 4k^2 + 1 - (2k)^2 = (2k^2 - 2k + 1)(2k^2 + 2k + 1).$$

Note that $2k^2 - 2k + 1 < n$ for $k > 0$. The problem is that $2k^2 + 2k + 1$ is not less than n , so we still have to work a little bit. Namely, we will choose k such that $2k^2 + 2k + 1$ is a multiple of 5, for instance choose $k = 5t + 1$, then

$$2k^2 + 2k + 1 = 5(10t^2 + 6t + 1).$$

Thus

$$n^2 + 1 = 5(10t^2 + 6t + 1)(2k^2 - 2k + 1)$$

and the numbers $5, 10t^2 + 6t + 1, 2k^2 - 2k + 1$ are pairwise distinct and less than n for $t \geq 1$. Thus their product divides $n!$. \square

Remark 8.1. Problem 11358 in AMM generalizes the previous result as follows: for any $d \geq 1$ there are infinitely many positive integers n such that $dn^2 + 1 \mid n!$. We leave it to the reader to check that for each $k \geq 2$ the number

$$n_k = dk^2(d+1)^2 + k(d+1) + 1$$

satisfies

$$dn_k^2 + 1 = (dk^2(d+1)^2 + 1)(d+1)(d^2k^2(d+1) + 2dk + 1)$$

and to deduce that $dn_k^2 + 1 \mid n_k!$ for all $k > 1$.

16. (Vietnam 2001) Let $(a_n)_{n \geq 1}$ be an increasing sequence of positive integers such that $a_{n+1} - a_n \leq 2001$ for all n . Prove that there are infinitely many pairs (i, j) with $i < j$ such that $a_i \mid a_j$.

Proof. Replace 2001 by an arbitrary positive integer k and imagine the following infinite matrix with k columns: the first row consists of the numbers $a_1 + 1, a_1 + 2, \dots, a_1 + k$. If the j th row is $x + 1, x + 2, \dots, x + k$, then the $j + 1$ th row is $N + x + 1, N + x + 2, \dots, N + x + k$, where

$$N = (x + 1)(x + 2) \dots (x + k)$$

is the product of the numbers on the j th row. Clearly if $a < b$ are on the same column then $a \mid b$. By assumption, among k consecutive positive integers greater than a_1 there is at least one term of the sequence, so each row of this matrix contains at least one term of the sequence. On the other hand, if we choose any $k + 1$ consecutive rows of the matrix, there will be at least two terms of the sequence in the same column (as there are at least $k + 1$ terms of the sequence in the corresponding sub-matrix, and only k columns). These two terms are distinct and one of them divides the other one. Since the $k + 1$ consecutive lines were arbitrary, it is clear that this procedure generates infinitely many pairs of distinct terms of the sequence in which one divides the other. \square

17. (Tournament of the Towns) Define a sequence $(a_n)_{n \geq 0}$ by $a_0 = 9$ and $a_{n+1} = a_n^3(3a_n + 4)$ for $n \geq 0$. Prove that $a_n + 1$ is a multiple of 10^{2^n} for all n .

Proof. We prove this by induction, the case $n = 0$ being clear. Assume now that $a_n + 1 = k \cdot 10^{2^n}$ for some integer k . A brutal expansion shows that

$$a_n^3 = (k \cdot 10^{2^n} - 1)^3 \equiv 3k \cdot 10^{2^n} - 1 \pmod{10^{2^{n+1}}}.$$

Therefore

$$a_{n+1} \equiv (3k \cdot 10^{2^n} - 1)(3k \cdot 10^{2^n} + 1) = 9k^2 \cdot 10^{2^{n+1}} - 1 \equiv -1 \pmod{10^{2^{n+1}}},$$

as needed.

We remark that the identity

$$x^3(3x + 4) + 1 = (x + 1)^2(3x^2 - 2x + 1),$$

which can be checked by a direct inspection of both sides, shows that $(a_n + 1)^2$ divides $a_{n+1} + 1$, yielding also the result immediately. \square

18. Find the largest integer k which divides $8^{n+1} - 7n - 8$ for all positive integers n .

Proof. Taking $n = 1$, we obtain $k \mid 49$ and so $k \leq 49$. We will prove that $49 \mid 8^{n+1} - 7n - 8$ for all n , which will show that the answer of the problem is 49. Using the binomial formula, we have

$$\begin{aligned} 8^{n+1} &= (1 + 7)^{n+1} = 1 + 7(n+1) + \binom{n+1}{2} 7^2 + \dots + 7^{n+1} \\ &\equiv 1 + 7(n+1) = 7n + 8 \pmod{49}, \end{aligned}$$

as desired. \square

19. Let a, b be distinct integers and let n be a positive integer. Prove that $(a - b)^2 \mid a^n - b^n$ if and only if $a - b \mid nb^{n-1}$.

Proof. Write $a - b = k$, then $(a - b)^2 \mid a^n - b^n$ if and only if $k^2 \mid (k + b)^n - b^n$. Using the binomial formula, we obtain

$$(k+b)^n - b^n = k^n + \binom{n}{1} k^{n-1} b + \dots + \binom{n}{n-1} k b^{n-1} \equiv n k b^{n-1} \pmod{k^2}.$$

Thus $k^2 \mid (k + b)^n - b^n$ if and only if $k^2 \mid n k b^{n-1}$, or equivalently (since $k \neq 0$) $k \mid n b^{n-1}$. \square

20. (BAMO 2012) Let n be a positive integer such that 81 divides both n and the number obtained by reversing the order of the digits of n . Prove that 81 also divides the sum of digits of n .

Proof. The binomial formula yields $10^k = (1+9)^k \equiv 1+9k \pmod{81}$ for all $k \geq 0$. Writing $n = a_0 + 10a_1 + \dots + 10^k a_k$ for the decimal expansion of n , we obtain

$$n \equiv \sum_{i=0}^k a_i(1+9i) = \sum_{i=0}^k a_i + 9 \sum_{i=0}^k i a_i \pmod{81}.$$

Let $n' = a_k + 10a_{k-1} + \dots + 10^k a_0$ be the number obtained by reversing the order of the digits of n . Then similarly

$$n' \equiv \sum_{i=0}^k a_i + 9 \sum_{i=0}^k (k-i)a_i.$$

Since n and n' are multiples of 81, so is $n + n'$, and using the previous congruences we deduce that

$$2 \sum_{i=0}^k a_i + 9k \sum_{i=0}^k a_i = (9k+2)S$$

is a multiple of 81, where S is the sum of digits of n . Thus $81 \mid (9k+2)S$ and it is an easy exercise left to the reader to deduce that S is a multiple of 9 (since $2S$ is a multiple of 9) and then that $81 \mid S$. \square

21. Prove that for all $n \geq 1$ the number $\frac{(2n)!(3n)!}{n!^5}$ is an integer multiple of $(n+1)^2$.

Proof. We have

$$\frac{(2n)!(3n)!}{(n+1)^2 n!^5} = \left(\frac{1}{n+1} \binom{2n}{n} \right)^2 \cdot \binom{3n}{n}$$

and $n+1 \mid \binom{2n}{n}$ by example 2.54, yielding the desired result. \square

22. Find all integers a such that n^2 divides $(n + a)^n - a$ for all positive integers n .

Proof. The binomial formula shows that

$$(n + a)^n - a \equiv a^n - a \pmod{n^2}.$$

Thus we must find a such that n^2 divides $a^n - a$ for all $n \geq 1$. Clearly $a = 0$ and $a = 1$ are solutions, while $a = -1$ is not (choose $n = 2$). Assume that $k = |a| > 1$ and choose $n = k$, so that $k^2 \mid a^k - a$. However $k^2 \nmid a^k$ (since $k > 1$), thus we must have $k^2 \mid a$ and then $k^2 \mid k$. This is however impossible for $k > 1$. Therefore the solutions of the problem are $a = 0$ and $a = 1$. \square

23. (P. Erdős) Prove that every positive integer is a sum of one or more numbers of the form $2^r \cdot 3^s$, where r and s are nonnegative integers and no summand divides another.

Proof. We proceed by induction, with base case $1 = 2^0 3^0$. Suppose all integers less than $n - 1$ can be represented. If n is even, then we can take a representation of $n/2$ and multiply each term by 2 to obtain a representation of n . If n is odd, take m so that $3^m \leq n < 3^{m+1}$. If $3^m = n$, we are done. Otherwise, choose a representation $(n - 3^m)/2 = s_1 + \dots + s_k$ in the desired form. Then $n = 3^m + 2s_1 + \dots + 2s_k$, and clearly none of the $2s_i$ divide each other or 3^m . Moreover, since $2s_i \leq n - 3^m < 3^{m+1} - 3^m$, we have $s_i < 3^m$, so 3^m cannot divide $2s_i$ either. Thus n has a representation of the desired form in all cases, completing the induction. Finally, note that the representations need not be unique: for instance, $11 = 2 + 3^2 = 3 + 2^3$. \square

24. (Kvant M 2274) Let $k \geq 2$ be an integer. Find all positive integers n such that 2^k divides $1^n + 2^n + \dots + (2^k - 1)^n$.

Proof. We will prove that the solutions of the problem are the odd numbers $n \geq 3$. Suppose first that n is odd, then

$$\begin{aligned} & 1^n + 2^n + \cdots + (2^k - 1)^n \\ &= (1^n + (2^k - 1)^n) + \cdots + ((2^{k-1} - 1)^n + (2^{k-1} + 1)^n) + (2^{k-1})^n \end{aligned}$$

and each term in the sum except for the last one is a multiple of 2^k (recall that $a^n + b^n$ is divisible by $a + b$ for all integers a, b). Thus the sum is a multiple of 2^k if and only if $(2^{k-1})^n$ is a multiple of 2^k , which happens if and only if $n \geq 3$.

Now let n be even. We shall prove by induction on k that

$$S_{n,k} := 1^n + 2^n + \cdots + (2^k - 1)^n$$

is not divisible by 2^k . This is true for $k = 2$ since

$$S_{n,2} = 1^n + 2^n + 3^n \equiv 2 \pmod{4}$$

when n is even. Suppose that 2^k does not divide $S_{n,k}$. Since

$$a^n \equiv (2^{k+1} - a)^n \pmod{2^{k+1}}$$

for all integers a , it follows that

$$S_{n,k+1} \equiv 2(1^n + 2^n + \cdots + (2^k - 1)^n) + 2^{kn} \equiv 2S_{n,k} \pmod{2^{k+1}},$$

which proves that $S_{n,k+1}$ is not divisible by 2^{k+1} . \square

25. Let k be an integer greater than 1 and let a_1, \dots, a_n be integers such that

$$a_1 + 2^i a_2 + 3^i a_3 + \cdots + n^i a_n = 0$$

for all $i = 1, 2, \dots, k-1$. Prove that $a_1 + 2^k a_2 + \cdots + n^k a_n$ is divisible by $k!$.

Proof. If b_0, b_1, \dots, b_{k-1} are integers, then

$$\begin{aligned} b_0(a_1 + 2a_2 + \dots + na_n) + b_1(a_1 + 2^2a_2 + \dots + n^2a_n) + \dots \\ + b_{k-1}(a_1 + 2^{k-1}a_2 + \dots + n^{k-1}a_n) = 0. \end{aligned}$$

We can rearrange this as

$$\begin{aligned} a_1(b_0 + b_1 + \dots + b_{k-1}) + a_2(2b_0 + 2^2b_1 + \dots + 2^{k-1}b_{k-1}) + \dots \\ + a_n(nb_0 + \dots + n^{k-1}b_n) = 0. \end{aligned}$$

It follows that for any polynomial $P(X) = b_0X + b_1X^2 + \dots + b_{k-1}X^{k-1}$ with integer coefficients, degree not exceeding $k-1$, and constant term 0 we have

$$a_1P(1) + a_2P(2) + \dots + a_nP(n) = 0.$$

The polynomial $P(X) = X^k - X(X-1)\dots(X-k+1)$ satisfies all previous conditions, and the previous relation can be written

$$a_1 + 2^k a_2 + \dots + n^k a_n = \sum_{i=1}^n a_i i(i-1)\dots(i-k+1) = k! \sum_{i=1}^n a_i \binom{i}{k}.$$

The right-hand side being a multiple of $k!$, we are done. \square

26. Prove that for any integer $k \geq 3$ there are k pairwise distinct positive integers such that their sum is divisible by each of the given numbers.

Proof. It suffices to prove the existence of pairwise distinct positive integers a_1, a_2, \dots, a_k such that

$$\frac{1}{a_1} + \frac{1}{a_2} + \dots + \frac{1}{a_k} = 1,$$

as then setting

$$b_1 = \frac{a_1 a_2 \dots a_k}{a_1}, \quad b_2 = \frac{a_1 a_2 \dots a_k}{a_2}, \dots, b_k = \frac{a_1 a_2 \dots a_k}{a_k}$$

yields the desired result. Let us now prove by induction the existence of a_1, \dots, a_k . For $k = 3$ choose $a_1 = 2, a_2 = 3$ and $a_3 = 6$. Assuming that a_1, \dots, a_k are pairwise distinct positive integers whose sum of inverses is 1, and $a_k = \max(a_1, \dots, a_k)$, the numbers $a_1, a_2, \dots, a_{k-1}, a_k + 1, a_k(a_k + 1)$ are pairwise distinct positive integers and the sum of their inverses is 1. \square

27. (Kvant) Prove that for any integer $n > 1$ there exist n pairwise distinct positive integers such that for any two a, b among them the number $a + b$ is divisible by $a - b$.

Proof. We prove this by induction on n . For $n = 2$ consider the numbers 1, 2. Assume that the result holds for n , thus there are integers $1 \leq a_1 < a_2 < \dots < a_n$ such that $a_i + a_j$ is divisible by $a_i - a_j$ for all $i \neq j$. Define

$$b_0 = a_1 a_2 \dots a_n \cdot \prod_{1 \leq i < j \leq n} (a_j - a_i)$$

and $b_i = a_i + b_0$ for $1 \leq i \leq n$. We will prove that b_0, b_1, \dots, b_n satisfy the desired properties. For all $1 \leq i \leq n$ we have $b_i - b_0 = a_i \mid b_i + b_0$ since a_i divides b_0 . Next, for $1 \leq i < j \leq n$ we have

$$b_j - b_i = a_j - a_i \mid a_i + a_j + 2b_0 = b_i + b_j,$$

since $a_i - a_j$ divides $a_i + a_j$ and $a_i - a_j$ divides b_0 . The result follows. \square

28. (Romania TST 1987) Let a, b, c be integers such that $a + b + c$ divides $a^2 + b^2 + c^2$. Prove that $a + b + c$ divides $a^n + b^n + c^n$ for infinitely many positive integers n .

Proof. Since $(a + b + c)^2 = a^2 + b^2 + c^2 + 2(ab + bc + ca)$, it follows that $a + b + c$ divides $2(ab + bc + ca)$. Next,

$$(a^2 + b^2 + c^2)^2 = a^4 + b^4 + c^4 + 2(a^2 b^2 + b^2 c^2 + c^2 a^2)$$

and

$$2(a^2 b^2 + b^2 c^2 + c^2 a^2) = 2(ab + bc + ca)(ab + bc + ca) - 4abc(a + b + c)$$

is a multiple of $a + b + c$. Thus $a + b + c$ divides $2(a^2b^2 + b^2c^2 + c^2a^2)$ and also $a^4 + b^4 + c^4$. We will prove by induction on n that $a + b + c$ divides $a^{2^n} + b^{2^n} + c^{2^n}$ and $2((ab)^{2^n} + (bc)^{2^n} + (ca)^{2^n})$ for $n \geq 1$. This has already been established for $n = 1$, so assume that it holds for n and let us prove it for $n + 1$. The proof is exactly as above, based on the identities

$$a^{2^{n+1}} + b^{2^{n+1}} + c^{2^{n+1}} = (a^{2^n} + b^{2^n} + c^{2^n})^2 - 2((ab)^{2^n} + (bc)^{2^n} + (ca)^{2^n})$$

and

$$\begin{aligned} & (ab)^{2^{n+1}} + (bc)^{2^{n+1}} + (ca)^{2^{n+1}} \\ &= ((ab)^{2^n} + (bc)^{2^n} + (ca)^{2^n})^2 - 2(abc)^{2^n}(a^{2^n} + b^{2^n} + c^{2^n}). \end{aligned}$$

We also present an alternate solution suggested by Richard Stong. Let $S = a + b + c$. Note that since $(a + b + c)^2 = a^2 + b^2 + c^2 + 2(ab + bc + ca)$, it follows that $S \mid 2(ab + bc + ca)$. Let $P_n = a^n + b^n + c^n$. Since a, b, c are the three roots of

$$(X - a)(X - b)(X - c) = X^3 - SX^2 + (ab + bc + ca)X - abc,$$

we see that

$$P_{n+3} = SP_{n+2} - (ab + bc + ca)P_{n+1} + abcP_n.$$

Note that by the hypotheses of the problem S divides P_1 and P_2 . We want to show S divides P_n for infinitely many n .

Now we consider two cases. If S is odd, then $S \mid ab + bc + ca$. Hence it follows from the recursion above that if S divides P_n , then S also divides P_{n+3} . Hence by a trivial induction S divides P_{3k+1} and P_{3k+2} for all $k \geq 0$. If S is even, then P_n is even for all n , hence S always divides $(ab + bc + ca)P_{n+1}$. We again conclude that if S divides P_n , then S divides P_{n+3} and hence S divides P_{3k+1} and P_{3k+2} for all $k \geq 0$. \square

29. (Russia 1995) Let a_1 be an integer greater than 1. Prove that there is an increasing sequence of positive integers $a_1 < a_2 < \dots$ such that

$$a_1 + a_2 + \dots + a_k \mid a_1^2 + \dots + a_k^2$$

for all $k \geq 1$.

Proof. We will construct such a sequence inductively. Assume that a_1, \dots, a_{k-1} have already been constructed and let us try to construct a_k . To simplify notations, let

$$x = a_1 + \dots + a_{k-1}, \quad y = a_1^2 + \dots + a_{k-1}^2.$$

We want to ensure that $a_k + x \mid a_k^2 + y$. Since $a_k + x \mid a_k^2 - x^2$, it suffices to ensure that

$$a_k + x \mid (a_k^2 + y) - (a_k^2 - x^2) = x^2 + y$$

and the easiest way to realize this is to take

$$a_k = x^2 + y - x = x(x - 1) + y.$$

Since $a_{k-1} > 1$ and $x \geq a_{k-1}$, $y \geq a_{k-1}^2$, it is clear that $a_k > a_{k-1}$. By construction, we have $a_1 + a_2 + \dots + a_k \mid a_1^2 + \dots + a_k^2$ and the result follows. \square

30. Let n be a positive integer. Prove that

- All multiples of $10^n - 1$ which do not exceed $10^n(10^n - 1)$ have sum of digits $9n$.
- The sum of digits of any multiple of $10^n - 1$ is at least $9n$.

Proof. a) Consider a multiple $N = (10^n - 1)k$ of $10^n - 1$ that does not exceed $10^n(10^n - 1)$, thus $k \leq 10^n$. Deleting the last zeros of k does not change the sum of digits of N , so we may assume that k is not a multiple of 10. In particular, $k < 10^n$ and so we can find digits a_0, \dots, a_{n-1} such that $k = a_0 + \dots + a_{n-1}10^{n-1}$ and $a_0 \neq 0$ (we do not impose $a_{n-1} \neq 0$). Now the subtraction algorithm or a direct algebraic manipulation show that

$$\begin{aligned} N &= (10^n - 1)k = \overline{a_{n-1} \dots a_0 00 \dots 0} - \overline{a_{n-1} \dots a_0} \\ &= \overline{a_{n-1} \dots a_1 (a_0 - 1) (9 - a_{n-1}) \dots (9 - a_1) (10 - a_0)}. \end{aligned}$$

The sum of digits of the last number is clearly $9n$.

b) Let $s(x)$ be the sum of digits of x . We will prove by strong induction on k that $s((10^n - 1)k) \geq 9n$ for all $k \geq 1$. For $k \leq 10^n$ this has already been seen in a). Assume that $k > 1$ and $s((10^n - 1)j) \geq 9n$ for $1 \leq j < k$ and write $(10^n - 1)k$ in base 10^n as

$$(10^n - 1)k = b_0 + b_1 \cdot 10^n + \dots + b_d \cdot (10^n)^d$$

for some $b_i \in \{0, 1, \dots, 10^n - 1\}$ with $b_d \neq 0$. Since $10^n - 1$ divides $10^{sn} - 1$ for all $s \geq 1$, the previous equality shows that $10^n - 1$ divides $b_0 + b_1 + \dots + b_d$. Note that $b_0 + \dots + b_d < b_0 + 10^n b_1 + \dots + 10^{nd} b_d$ unless $d = 0$, but then $k = 1$, contradiction. So we can write $b_0 + b_1 + \dots + b_d = j(10^n - 1)$ for some $1 \leq j < k$. Now, since $b_i < 10^n$, we obtain

$$s(k(10^n - 1)) = s(b_0) + s(b_1) + \dots + s(b_d) \geq s(b_0 + \dots + b_d)$$

and by the inductive hypothesis the last quantity is greater than or equal to $9n$. The result follows. \square

Remark 8.2. We have freely used the inequality

$$s(a + b) \leq s(a) + s(b)$$

in the previous solution. We invite the reader to supply a proof.

31. (USAMO 1998) Prove that for each $n \geq 2$ there is a set S of n integers such that $(a - b)^2$ divides ab for every distinct $a, b \in S$.

Proof. We will construct such a set, consisting of nonzero integers, by induction. Take for $n = 2$ the set $\{1, 2\}$. Assume that such a set $S = \{a_1, \dots, a_n\}$ has been constructed. The new set T will be taken of the form

$$T = \{a_1 + k, \dots, a_n + k\} \cup \{k\}$$

for a suitable integer k .

We need $(a_i - a_j)^2 \mid (a_i + k)(a_j + k)$ and $a_i^2 \mid k(a_i + k)$ for all $i \neq j$ between 1 and n . The divisibility $a_i^2 \mid k(a_i + k)$ certainly holds if we impose $a_i^2 \mid k$ for all i (even $a_i \mid k$ would suffice). On the other hand,

since $(a_i - a_j)^2 \mid a_i a_j$, the divisibility $(a_i - a_j)^2 \mid (a_i + k)(a_j + k)$ holds if we impose $(a_i - a_j)^2 \mid k$ for all $i \neq j$. Thus it suffices to take any nonzero integer k which is a multiple of $\prod_{i=1}^n a_i^2 \cdot \prod_{1 \leq i < j \leq n} (a_i - a_j)^2$. \square

32. (Romania JBMO TST 2004) Let A be a set of positive integers such that

- a) if $a \in A$, then all positive divisors of a are also in A ;
- b) if $a, b \in A$ satisfy $1 < a < b$, then $1 + ab \in A$.

Prove that if A has at least 3 elements, then A is the set of all positive integers.

Proof. We start by proving that A contains 1, 2, 3, 4, 5. It is clear that $1 \in A$. If $2 \notin A$, then by a) all elements of A are odd. Since A has at least three elements, we can choose $a, b \in A$ with $1 < a < b$. By b), $1 + ab \in A$, but $1 + ab$ is clearly even, a contradiction. Hence $2 \in A$.

Next, we prove that A contains a multiple of 4, and hence $4 \in A$. Choose any $a > 2$ in A (possible, since $|A| \geq 3$). Then applying successively property b) we obtain $1 + 2a \in A$, then $1 + 2(1 + 2a) = 3 + 4a \in A$ and finally $b = 1 + (1 + 2a)(3 + 4a) \in A$. Note that $b > 2$ is even. Applying the same argument, $c = 1 + (1 + 2b)(3 + 4b) \in A$, but this last number is a multiple of 4, hence we are done. It also follows that $1 + 2 \cdot 4 = 9 \in A$, hence $3 \in A$, then $1 + 2 \cdot 3 = 7 \in A$, $1 + 2 \cdot 7 = 15 \in A$ and $5 \in A$. Also, $1 + 5 \cdot 7 = 36 \in A$, hence $6 \in A$.

It is now time to conclude: we will prove by strong induction on n that $n \in A$. By the previous work, we may assume that $n \geq 7$ and that $1, 2, \dots, n-1 \in A$. If n is odd, say $n = 2k+1$ for some $k > 2$, then $n \in A$ by property b), since $2, k \in A$. So assume that $n = 2k$ is even, with $k > 3$. Again, since $k, k-1 \in A$ are greater than 2, we have $1 + 2k \in A$ and $1 + 2(k-1) = 2k-1 \in A$. But then $1 + (2k-1)(2k+1) = 4k^2 \in A$, hence $n = 2k \in A$. The inductive step is proved and the result follows. \square

33. (USAMO 2002) Let a, b be integers greater than 2. Prove that there exists a positive integer k and a finite sequence n_1, n_2, \dots, n_k of positive

integers such that $n_1 = a$, $n_k = b$, and $n_i n_{i+1}$ is divisible by $n_i + n_{i+1}$ for each i ($1 \leq i < k$).

Proof. If a, b are positive integers, say that they are linked if there is a positive integer k and a finite sequence n_1, n_2, \dots, n_k of positive integers such that $n_1 = a$, $n_k = b$, and $n_i n_{i+1}$ is divisible by $n_i + n_{i+1}$ for each i ($1 \leq i < k$). It is clear that if a is linked to b and b is linked to c , then a is linked to c . Next, if $a > 1$ is odd, then a is linked to $a + 1$, since we can use the sequence $a, a^2 - a, a^2 + a, a + 1$. Also, if $a > 2$ is even, write $a = 2k$ and use the sequence $a, 2k^2 - 2k, 2k^2 + 2k, 2k + 2 = a + 2$ to link a and $a + 2$. We deduce that all even numbers are linked, and since any odd a is linked to the even number $a + 1$, it follows that all numbers greater than 2 are linked. \square

Remark 8.3. We suggest the reader to try to solve the following very similar problem (proposed in an Iranian Mathematical Olympiad in 2006): let m, n be integers greater than 2. Prove that there is a sequence a_0, \dots, a_k of integers greater than 1 such that $a_0 = m$, $a_k = n$ and $a_i + a_{i+1} \mid a_i a_{i+1} + 1$ for all $0 \leq i < k$.

34. Is it true that for any integer $k > 1$ we can find an integer $n > 1$ such that k divides each of the numbers $\binom{n}{1}, \binom{n}{2}, \dots, \binom{n}{n-1}$?

Proof. The answer is negative. We will show that for $k = 4$ there is no such n . Assume by contradiction that 4 divides each of the numbers $\binom{n}{1}, \binom{n}{2}, \dots, \binom{n}{n-1}$. Then 4 also divides their sum, which is $2^n - 2$. This can only happen if $n = 1$ (as if $n > 1$ the number 2^n is a multiple of 4), however in this case $\binom{n}{1} = 1$ is not a multiple of 4. \square

35. (Catalan) Prove that $m!n!(m+n)!$ divides $(2m)!(2n)!$ for all positive integers m, n .

Proof. Let

$$f(m, n) = \frac{(2m)!(2n)!}{m!n!(m+n)!}.$$

We will prove by induction on m the following statement: for all $n \geq 1$ we have $f(m, n) \in \mathbf{Z}$. The case $m = 1$ follows directly from exercise 2.54. Assume now that the result holds for m and let us prove it for $m + 1$. Fix $n > 1$. Then direct computations yield

$$\begin{aligned} f(m, n-1) &= \frac{(2m)!(2n-2)!}{m!(n-1)!(m+n-1)!} \\ &= f(m, n) \cdot \frac{n(m+n)}{2n(2n-1)} = f(m, n) \cdot \frac{m+n}{2(2n-1)} \end{aligned}$$

and

$$\begin{aligned} f(m+1, n-1) &= \frac{(2m+2)!(2n-2)!}{(m+1)!(n-1)!(m+n)!} \\ &= f(m, n) \cdot \frac{(2m+1)(2m+2)n}{2n(2n-1)(m+1)} = \frac{2m+1}{2n-1} f(m, n). \end{aligned}$$

We deduce that

$$f(m+1, n-1) = 4f(m, n-1) - f(m, n).$$

The right-hand side is an integer by the inductive hypothesis. Thus $f(m+1, n-1)$ is an integer for all $n > 1$, which proves the inductive step and finishes the solution. \square

Remark 8.4. The previous solution is not natural and not easy to come up with, but with the tools we have developed so far it is not easy to find a natural solution for the previous problem. Once the theory of prime numbers and p -adic valuations is established (and this will occupy us quite some time in the book!), this problem will become a straightforward exercise.

36. Let $x_1 < x_2 < \dots < x_{n-1}$ be consecutive positive integers such that $x_k \mid k \binom{n}{k}$ for all $1 \leq k \leq n-1$. Prove that x_1 equals 1 or 2.

Proof. Let $x = x_1 - 1$ and assume that $x > 1$, i.e. that the conclusion fails. Note that $x_i = x + i$ for $1 \leq i \leq n-1$. The key ingredient is the

following identity

$$\frac{n!}{(x+1)\dots(x+n)} = \sum_{k=1}^n (-1)^{k-1} \frac{k \binom{n}{k}}{x+k}.$$

Let us take this for granted for a moment and see how to conclude. By assumption all terms but the last one in the above sum are integers. We deduce that

$$a := \frac{n!}{(x+1)\dots(x+n)} + (-1)^n \frac{n}{x+n}$$

is an integer. However, since $x \geq 2$ we have

$$|a| < \frac{n!}{2 \cdot \dots \cdot (n+1)} + \frac{n}{n+1} = 1,$$

thus $a = 0$. This already shows that n is odd, and also that

$$(x+1)\dots(x+n-1) = (n-1)!.$$

This is clearly impossible, since the left-hand side is greater than $(n-1)!$. Thus the problem is solved, once the identity is proved.

Let us prove now the identity. Multiplying by $(x+1)\dots(x+n)$, we are reduced to proving the identity

$$\begin{aligned} (x+2)\dots(x+n) \binom{n}{1} - (x+1)(x+3)\dots(x+n) 2 \binom{n}{2} + \dots \\ + (-1)^{n-1} n \binom{n}{n} (x+1)\dots(x+n-1) = n!. \end{aligned}$$

The difference $f(x)$ between the left-hand side and the right-hand side is a polynomial of degree at most $n-1$ in x , and one immediately checks that $f(-1) = f(-2) = \dots = f(-n) = 0$ (note that the complicated sum in the left-hand side has only one nonzero term when x is one of the numbers $-1, -2, \dots, -n$). Therefore the polynomial f is the zero polynomial, which finishes the proof of the identity. \square

37. Prove that for any $n > 1$ there are $2n - 2$ positive integers such that the average of any n of them is not an integer.

Proof. Choose arbitrary positive integers a_1, \dots, a_{n-1} which are divisible by n and arbitrary positive integers b_1, \dots, b_{n-1} congruent to 1 modulo n . It is clear that the numbers $a_1, \dots, a_{n-1}, b_1, \dots, b_{n-1}$ have the property that the average of any n of them is not an integer, since the sum of the n numbers gives a remainder between 1 and $n - 1$ when divided by n . \square

38. Let n be a positive integer. Find the remainder of 3^{2^n} when divided by 2^{n+3} .

Proof. We have

$$3^{2^n} - 1 = (3-1)(3+1)(3^2+1)\dots(3^{2^{n-1}}+1) = 8(3^2+1)(3^4+1)\dots(3^{2^{n-1}}+1).$$

Each of the numbers $3^2 + 1, 3^4 + 1, \dots, 3^{2^{n-1}} + 1$ is even and not divisible by 4, thus their product is of the form $2^{n-1}(2k + 1)$ for some $k > 0$. Then

$$3^{2^n} - 1 = 2^{n+2}(2k + 1) = 2^{n+3}k + 2^{n+2}$$

and so the required remainder is $2^{n+2} + 1$. \square

39. (Saint Petersburg 1996) Let P be a polynomial with integer coefficients, of degree greater than 1. Prove that there is an infinite arithmetic progression none of whose terms belongs to $\{P(n) \mid n \in \mathbf{Z}\}$.

Proof. Since $\deg P > 1$, the polynomial $P(X+1) - P(X)$ is not constant, thus we can find $x > 1$ such that the number $d = |P(x+1) - P(x)|$ satisfies $d > 1$. Since $P(x)$ and $P(x+1)$ give the same remainder when divided by d , there is r between 0 and $d-1$ such that none of the numbers $P(x), P(x+1), \dots, P(x+d-1)$ gives remainder r when divided by d . If m is any integer, we can find $y \in \{x, x+1, \dots, x+d-1\}$ such that $m \equiv y \pmod{d}$. Then $P(m) \equiv P(y) \pmod{d}$ and so the remainder of $P(m)$

when divided by d is not r . It follows that $\{P(n) \mid n \in \mathbf{Z}\}$ has empty intersection with the infinite arithmetic progression $r + d\mathbf{Z}$ consisting of numbers congruent to r modulo d . \square

40. (Baltic Way 2011) Determine all positive integers d such that whenever d divides a positive integer n , d also divides any integer obtained by rearranging the digits of n .

Proof. Let d be a solution of the problem. Choose a large integer N such that $10^N > n$. Among the consecutive integers

$$10^{N+1} + 2 \cdot 10^N, 10^{N+1} + 2 \cdot 10^N + 1, \dots, 10^{N+1} + 2 \cdot 10^N + 10^N - 1$$

there is a multiple of n . Such a number is of the form $\overline{12a_1 \dots a_n}$ for some digits a_1, \dots, a_n . By assumption d divides any number obtained by permuting the digits of $\overline{12a_1 \dots a_n}$, in particular it divides $\overline{a_1 \dots a_n 21}$ and $\overline{a_1 \dots a_n 12}$. Therefore d also divides the difference of these two numbers, which is 9. It follows that $d = 1, 3$ or 9 . Conversely, any divisor d of 9 is a solution of the problem. Indeed, assume that $d \mid n$ and that n' is obtained from n by permuting its digits. Then n' and n have the same sum of digits, say k . Since $n \equiv k \pmod{9}$ and $n' \equiv k \pmod{9}$, we have $n \equiv n' \pmod{9}$ and so $n \equiv n' \pmod{d}$, yielding $d \mid n'$. Thus the solutions of the problem are 1, 3, 9. \square

41. (Russia) A convex polygon on the coordinate plane contains at least $m^2 + 1$ points with integer coordinates in its interior. Show that some $m + 1$ of these points lie on a line.

Proof. For each point P with integer coefficients inside the polygon, consider the pair of remainders obtained by dividing the coordinates of P by m . We have at least $m^2 + 1$ pairs associated to the points with integer coordinates inside the polygon. On the other hand, since there are m remainders mod m , there are m^2 pairs of remainders mod m . Thus we can find two points P with coordinates a, b and Q with coordinates c, d such that $a \equiv c \pmod{m}$ and $b \equiv d \pmod{m}$. Then the points A_k

with coordinates $c + \frac{k}{m}(a - c)$ and $d + \frac{k}{m}(b - d)$, for $0 \leq k \leq m$, are on the segment with endpoints P, Q , have integer coordinates and are inside the polygon (since the polygon is convex). \square

42. (IMO 2001) Let $n > 1$ be an odd integer and let c_1, c_2, \dots, c_n be integers. For each permutation $a = a_1, a_2, \dots, a_n$ of $1, 2, \dots, n$, define

$$S(a) = c_1 a_1 + c_2 a_2 + \dots + c_n a_n.$$

Prove that there are permutations $a \neq b$ of $1, 2, \dots, n$ such that $n! \mid S(a) - S(b)$.

Proof. Suppose that for all permutations a and b of $1, 2, \dots, n$ the number $n!$ does not divide $S(a) - S(b)$. Since there are $n!$ remainders modulo $n!$ as well as $n!$ permutations of $1, 2, \dots, n$, it follows that the remainders of the numbers $S(a)$ (over all permutations a) when divided by $n!$ are $0, 1, \dots, n! - 1$ in some order, thus

$$\sum_a S(a) \equiv 1 + 2 + \dots + (n! - 1) = \frac{n!(n! - 1)}{2} \pmod{n!}.$$

On the other hand,

$$\sum_a S(a) = \sum_a \sum_{j=1}^n a_j c_j = \sum_{j=1}^n c_j \cdot \sum_a a_j.$$

For each $k \in \{1, 2, \dots, n\}$ there are precisely $(n-1)!$ permutations a for which $a_j = k$, thus

$$\sum_a a_j = \sum_{k=1}^n (n-1)!k = (n-1)! \cdot \frac{n(n+1)}{2} = n! \cdot \frac{n+1}{2} \equiv 0 \pmod{n!},$$

the last congruence uses the hypothesis that n is odd. Combining these congruences, we deduce that $n!$ divides $\frac{n!(n!-1)}{2}$, which is clearly absurd, since $n! - 1$ is odd. Hence our assumption was wrong and the result follows. \square

43. Let $n, k > 1$ be integers. Consider a set A of k integers. For each nonempty subset B of A , compute the remainder of the sum of elements of B when divided by n . Assume that 0 does not appear among these remainders. Prove that there are at least k distinct remainders obtained in this way. Moreover, if there are only k such remainders, then all elements of A give the same remainder when divided by n .

Proof. Let a_1, \dots, a_k be the elements of A . We claim that $a_1, a_1 + a_2, \dots, a_1 + \dots + a_k$ give pairwise distinct remainders when divided by n , which is enough to conclude for the first part of the problem. Indeed, if $a_1 + \dots + a_i$ and $a_1 + \dots + a_j$ give the same remainder for some $1 \leq i < j \leq k$, then $a_{i+1} + \dots + a_j$ is a multiple of n , contradicting the hypothesis.

Assume now that there are exactly k remainders, which must be the remainders of $a_1, a_1 + a_2, \dots, a_1 + \dots + a_k$. Assume that a_1 and a_2 give different remainders when divided by n . Thus there is $i \geq 2$ such that

$$a_2 \equiv a_1 + a_2 + \dots + a_i \pmod{n},$$

meaning that $a_1 + a_3 + \dots + a_i$ is a multiple of n , a contradiction. Thus $a_1 \equiv a_2 \pmod{n}$. But since the order of a_1, \dots, a_k is not relevant in the previous argument, we deduce that any two a_i 's are congruent mod n , and the problem is solved. \square

44. (IMO 2005) A sequence a_1, a_2, \dots of integers has the following properties:
- a) a_1, a_2, \dots, a_n is a complete residue system modulo n for all $n \geq 1$.
 - b) there are infinitely many positive and infinitely many negative terms in the sequence.

Prove that each integer appears exactly once in this sequence.

Proof. It is clear that each integer appears at most once, for if $a_m = a_n$ for some $m < n$, then a_1, \dots, a_n cannot be a complete residue system modulo n . Hence it remains to prove that each integer k actually appears in the sequence. By considering the sequence $a_1 - k, a_2 - k, \dots$ (which

satisfies the same properties as the original sequence), we reduce to the case $k = 0$.

Assume now that a_n is nonzero for all n . Replacing a_n by $-a_n$ for all n , properties a) and b) are still satisfied, so we may assume that $a_1 > 0$. Let n be the smallest positive integer for which $a_n < 0$ and let $i \in \{1, \dots, n-1\}$ be such that $a_i = \max(a_1, \dots, a_{n-1})$. Note that $a_i \geq n-1$, since a_1, \dots, a_{n-1} are pairwise distinct positive integers. Hence $N = a_i - a_n \geq n$. Since $a_i \equiv a_n \pmod{N}$, it follows that a_1, \dots, a_N cannot be a complete residue system modulo N , a contradiction. Hence $a_n = 0$ for some n and, as explained in the first paragraph, we are done.

Here is an alternate solution, due to Richard Stong. We will prove by induction on n that any sequence a_1, a_2, \dots satisfying condition (a) has the property that for all n the numbers a_1, \dots, a_n are consecutive integers in some order. Then from condition (b) the requested conclusion is almost immediate: by (b), the sequence contains arbitrarily large magnitude positive and negative integers, and since it has blocks of consecutive integers it must contain every integer in between.

For the inductive proof, the base case $n = 1$ is trivial. For the inductive step, suppose a_1, \dots, a_n are consecutive. That is, they are the numbers $a_i, a_i + 1, \dots, a_i + n - 1 = a_j$ for some $1 \leq i, j \leq n$. Clearly, a_{n+1} cannot be a repeat of one of these n numbers, otherwise a_1, \dots, a_{n+1} would not be a complete residue system mod $n + 1$. If $a_{n+1} > a_i + n$, then let $N = a_{n+1} - a_i \geq n + 1$. Since $a_{n+1} \equiv a_i \pmod{N}$, we see that a_1, \dots, a_N is not a complete residue system modulo N , a contradiction. Similarly, if $a_{n+1} < a_i - 1$, then we let $N = a_j - a_{n+1} \geq n + 1$ and $a_{n+1} \equiv a_j \pmod{N}$ gives a contradiction. Thus a_{n+1} must be either $a_i - 1$ or $a_i + n = a_j + 1$. In either case we see that a_1, \dots, a_{n+1} are $n + 1$ consecutive integers. \square

45. For a positive integer n , consider the set

$$S = \{0, 1, 1 + 2, 1 + 2 + 3, \dots, 1 + 2 + 3 + \dots + (n - 1)\}$$

Prove S is a complete residue system modulo n if and only if n is a power of 2.

Proof. First, assume that n is a power of 2, say $n = 2^k$. We need to prove that if $0 \leq i < j \leq n - 1$ satisfy $\frac{i(i+1)}{2} \equiv \frac{j(j+1)}{2} \pmod{n}$, then $i = j$. Note that

$$\frac{i(i+1)}{2} - \frac{j(j+1)}{2} = \frac{i^2 - j^2 + i - j}{2} = \frac{(i-j)(i+j+1)}{2}.$$

So, assume that 2^{k+1} divides $(i-j)(i+j+1)$. One of the numbers $i-j$ and $i+j+1$ is odd, hence 2^{k+1} divides either $j-i$ or $i+j+1$. Since both these numbers have absolute value less than 2^{k+1} , this is only possible when one of them is 0, that is $i = j$.

Next, assume that n is not a power of 2 and write $n = 2^k m$ with $k \geq 0$ and $m > 1$ odd. Choose an integer $j \in \{0, 1, \dots, m-1\}$ such that $m \mid 2j+1+2^{k+1}$ (this is possible since m is odd) and set $i = j + 2^{k+1}$. Then $i \in \{0, 1, \dots, n-1\}$, n does not divide $i-j = 2^{k+1}$ and yet n divides

$$\frac{i(i+1)}{2} - \frac{j(j+1)}{2} = \frac{(i-j)(i+j+1)}{2},$$

since 2^k divides $\frac{i-j}{2}$ and m divides $i+j+1$ by construction. Thus S is not a complete residue system modulo n , a contradiction. \square

46. (Argentina 2008) 101 positive integers are written on a line. Prove that we can write signs $+$, signs \times and parentheses between them, without changing the order of the numbers, in such a way that the resulting expression makes sense and the result is divisible by $16!$.

Proof. By example 2.89 for any integers a_1, \dots, a_m we can find $i < j$ such that m divides $a_{i+1} + \dots + a_j$. In particular

$$m \mid (a_1 + \dots + a_i) \times (a_{i+1} + \dots + a_j) \times (a_{j+1} + \dots + a_m).$$

We deduce that if a, b are positive integers and $m = ab$, $n = a + b$, then for any sequence of n integers we can insert parentheses and signs $+$, \times around the first a terms to make the result divisible by a , and around the last b terms to make the result divisible by b , and finally enclose these

two within parentheses and add a multiplication operation to make the result divisible by m . The result follows by observing that

$$m = 16! = 2^{15} \times 3^6 \times 5^3 \times 7^2 \times 11 \times 13$$

and

$$30 + 18 + 15 + 14 + 11 + 13 = 101. \quad \square$$

47. (adapted from Kvant M33) Consider the remainders of 2^n when divided by $1, 2, \dots, n$. Prove that their sum exceeds $cn \log n$ for some constant $c > 0$ (independent of $n > 1$).

Proof. If $k > 1$ is odd, then the remainder of 2^n when divided by $2^i k$ is divisible by 2^i and nonzero, hence it must be at least 2^i . Let x_i be the number of positive integers of the form $2^i(2k+1)$ with $k \geq 1$ and $2^i(2k+1) \leq n$. Then clearly

$$x_i = \left\lfloor \frac{n - 2^i}{2^{i+1}} \right\rfloor \geq \frac{n - 3 \cdot 2^i}{2^{i+1}}$$

for all i . If N is chosen such that $3 \cdot 2^N \leq n < 3 \cdot 2^{N+1}$, the previous observations show that

$$\begin{aligned} \sum_{k=1}^n (2^n \pmod k) &\geq x_0 + 2x_1 + \dots + 2^N x_N \geq \sum_{i=0}^N \frac{n - 3 \cdot 2^i}{2} \\ &= (N+1) \frac{n}{2} - \frac{3}{2} (2^{N+1} - 1). \end{aligned}$$

Using the inequalities $3 \cdot 2^N \leq n < 3 \cdot 2^{N+1}$, it is easy to see that the last expression exceeds $\frac{n}{2}(\log_2(n) - 4)$, which yields the desired result. \square

8.2 GCD and LCM

1. Prove that for all positive integers a, b, c we have

$$\gcd(a, bc) \mid \gcd(a, b) \cdot \gcd(a, c).$$

Proof. Let $d = \gcd(a, b)$ and write $a = du$ and $b = dv$ with $\gcd(u, v) = 1$. We need to prove that $d \gcd(u, vc) \mid d \gcd(a, c)$, or equivalently $\gcd(u, vc) \mid \gcd(a, c)$. But $\gcd(u, vc)$ divides u , so it is relatively prime to v (since $\gcd(u, v) = 1$). Since $\gcd(u, vc)$ also divides vc , Gauss' lemma yields $\gcd(u, vc) \mid c$. Since it is clear that $\gcd(u, vc) \mid u \mid a$, the result follows. \square

2. (Romania TST 1990) Let a, b be relatively prime positive integers. Let x, y be nonnegative integers and let n be a positive integer for which

$$ax + by = a^n + b^n.$$

Prove that

$$\left\lfloor \frac{x}{b} \right\rfloor + \left\lfloor \frac{y}{a} \right\rfloor = \left\lfloor \frac{a^{n-1}}{b} \right\rfloor + \left\lfloor \frac{b^{n-1}}{a} \right\rfloor.$$

Proof. Reducing the first equation modulo a and b and using the fact that $\gcd(a, b) = 1$ we obtain $y \equiv b^{n-1} \pmod{a}$ and $x \equiv a^{n-1} \pmod{b}$. Thus we can find integers c, d such that $y = b^{n-1} + ca$ and $x = a^{n-1} + db$. Replacing these relations in the equation $ax + by = a^n + b^n$, we obtain $c + d = 0$. But then

$$\left\lfloor \frac{x}{b} \right\rfloor + \left\lfloor \frac{y}{a} \right\rfloor = \left\lfloor \frac{a^{n-1}}{b} + d \right\rfloor + \left\lfloor \frac{b^{n-1}}{a} + c \right\rfloor$$

and the result follows from the 1-periodicity of the floor function. \square

3. (Kvant M 1996) Find all $n > 1$ for which there exist pairwise different positive integers a_1, a_2, \dots, a_n such that

$$\frac{a_1}{a_2} + \frac{a_2}{a_3} + \dots + \frac{a_{n-1}}{a_n} + \frac{a_n}{a_1}$$

is an integer.

Proof. For every $n \geq 3$ consider the positive integers $a_1 = 1$, $a_2 = n - 1, \dots, a_n = (n - 1)^{n-1}$. Then

$$\frac{a_1}{a_2} + \frac{a_2}{a_3} + \dots + \frac{a_{n-1}}{a_n} + \frac{a_n}{a_1} = \frac{1}{n-1} + \frac{n-1}{(n-1)^2} + \dots + \frac{(n-1)^{n-2}}{(n-1)^{n-1}} + \frac{(n-1)^{n-1}}{1}$$

is an integer, equal to $1 + (n - 1)^{n-1}$. Suppose now that $a_1 \neq a_2$ and $\frac{a_1}{a_2} + \frac{a_2}{a_1}$ is an integer. Dividing a_1 and a_2 by their gcd, we may assume that they are relatively prime.

Then $a_1 a_2 \mid a_1^2 + a_2^2$ and $\gcd(a_1, a_1^2 + a_2^2) = \gcd(a_1, a_2^2) = 1$, thus necessarily $a_1 = 1$ and similarly $a_2 = 1$, a contradiction. \square

4. Let m, n be positive integers greater than 1.

We define the sets $P_m = \left\{ \frac{1}{m}, \frac{2}{m}, \dots, \frac{m-1}{m} \right\}$ and $P_n = \left\{ \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n} \right\}$. Find

$$\min\{|a - b| : a \in P_m, b \in P_n\}$$

Proof. We need to find the smallest value that $f(i, j) := \left| \frac{i}{m} - \frac{j}{n} \right|$ can take when $1 \leq i < m$ and $1 \leq j < n$. If $\gcd(m, n) = d > 1$, then we can take $i = \frac{m}{d}$ and $j = \frac{n}{d}$ and get $f(i, j) = 0$, thus the answer of the problem is 0 if $\gcd(m, n) > 1$.

Assume now that $\gcd(m, n) = 1$. We cannot have $f(i, j) = 0$, since if $in = jm$ then $m \mid in$ and $m \mid i$ (since $\gcd(m, n) = 1$), contradicting the inequalities $1 \leq i < m$. Thus

$$f(i, j) = \frac{|in - jm|}{mn} \geq \frac{1}{mn}.$$

We will prove that we can find i, j such that $|in - jm| = 1$, which will imply that the answer of the problem is $\frac{1}{mn}$ when $\gcd(m, n) = 1$. Since $n, 2n, \dots, (m - 1)n$ give pairwise distinct and nonzero remainders when divided by m , one of them say in gives remainder 1 and so $in = 1 + jm$ for some integer j . Since $1 \leq i < m$, we have $1 \leq j < n$ and the problem is solved. \square

5. (Saint Petersburg 2004) Positive integers m, n, k are such that $5^n - 2$ and $2^k - 5$ are multiples of $5^m - 2^m$. Prove that $\gcd(m, n) = 1$.

Proof. Let $d = \gcd(m, n)$. Then $5^d - 2^d \mid 5^m - 2^m$ and $5^d - 2^d \mid 5^{kn} - 2^{kn}$. But

$$5^{kn} - 2^{kn} \equiv (5^n)^k - (2^k)^n \equiv 2^k - 5^n \equiv 5 - 2 = 3 \pmod{5^m - 2^m}.$$

It follows that $5^d - 2^d \mid 3$ and so $d = 1$. □

6. (Russia 2000) Sasha tries to find a positive integer $X \leq 100$. He can choose any two positive integers M, N less than 100 and ask for $\gcd(X + M, N)$. Prove that he can find X after 7 questions.

Proof. Let $f(n)$ be the remainder of X modulo 2^n . Since $X \leq 100$, we have $X = f(7)$. Note that $f(n+1) = f(n)$ or $f(n+1) = f(n) + 2^n$, the last equality happening if and only if $\gcd(X + 2^n - f(n), 2^{n+1}) = 2^n$. Thus Sasha can find $f(6)$ after 6 questions, since he knows $f(0) = 1$ and the previous discussion shows that if he knows $f(n)$, then he also knows $f(n+1)$, as long as $n+1 \leq 6$ (to ensure that $2^{n+1} < 100$). Thus after 6 questions Sasha knows that X is either $f(6)$ or $f(6) + 64$. His final question will be to compute $\gcd(X + M, 3)$, where $M \in \{1, 2, 3\}$ is chosen such that $3 \mid f(6) + M$. If he gets the answer 3, then $X = f(6)$, otherwise $X = f(6) + 64$. Hence after this new question Sasha knows X . □

7. (Poland 2002) Let k be a fixed positive integer. The sequence $\{a_n\}_{n \geq 1}$ is defined by

$$a_1 = k + 1, a_{n+1} = a_n^2 - ka_n + k.$$

Show that if $m \neq n$, then the numbers a_m and a_n are relatively prime.

Proof. Write the recurrence relation as

$$a_{n+1} - k = a_n(a_n - k).$$

An immediate induction using this relation shows that $a_n > k$ for all k and also that $a_n \equiv 1 \pmod{k}$ for all n . Next, multiply the previous relations to get

$$\prod_{i=1}^{n-1} (a_{i+1} - k) = \prod_{i=1}^{n-1} a_i \cdot \prod_{i=1}^{n-1} (a_i - k),$$

which, after division by $\prod_{i=1}^{n-1} (a_i - k)$, can be written as

$$a_n - k = a_1 a_2 \dots a_{n-1}.$$

Now, if d divides a_n and a_m for some $m < n$, then it divides $a_1 a_2 \dots a_{n-1} = a_n - k$ and a_n , thus it also divides k and a_n . Since $a_n \equiv 1 \pmod{k}$, it follows that d divides both a_n and $a_n - 1$ and so $d = 1$. \square

8. (Romania TST 2005) Let m, n be relatively prime positive integers with m even and n odd. Prove that

$$\sum_{k=1}^{n-1} (-1)^{\lfloor \frac{mk}{n} \rfloor} \left\{ \frac{mk}{n} \right\} = \frac{1}{2} - \frac{1}{2n}.$$

We denote by $\{x\}$ the fractional part of x , i.e. $\{x\} = x - \lfloor x \rfloor$.

Proof. Write the Euclidean division of mk by n as $mk = q_k n + r_k$ for $1 \leq k \leq n-1$. Since $\gcd(m, n) = 1$, the remainders r_1, \dots, r_{n-1} are pairwise distinct and nonzero, thus they must be a permutation of $1, 2, \dots, n-1$.

On the other hand we have

$$\left\lfloor \frac{mk}{n} \right\rfloor = q_k, \quad \left\{ \frac{mk}{n} \right\} = \frac{mk}{n} - q_k = \frac{r_k}{n}.$$

Thus the equality is equivalent to

$$\sum_{k=1}^{n-1} (-1)^{q_k} r_k = \frac{n-1}{2}.$$

Now, since m is even and n is odd, we have

$$0 \equiv mk = q_k n + r_k \equiv q_k + r_k \pmod{2},$$

thus $(-1)^{q_k} = (-1)^{r_k}$ and we are reduced to proving that

$$\sum_{k=1}^{n-1} (-1)^{r_k} r_k = \frac{n-1}{2}.$$

Taking into account the first paragraph, this is equivalent to

$$\sum_{k=1}^{n-1} (-1)^k k = \frac{n-1}{2},$$

which follows immediately by induction on n (going from n to $n+2$, since n is assumed to be odd). \square

9. An infinite sequence a_1, a_2, \dots of positive integers has the property that $\gcd(a_m, a_n) = \gcd(m, n)$ for all $m \neq n \geq 1$. Prove that $a_n = n$ for all $n \geq 1$.

Proof. Taking $m = 2n$ yields $\gcd(a_{2n}, a_n) = n$, thus $n \mid a_n$. Suppose that $a_n \neq n$ for some n . Then $\gcd(a_{a_n}, a_n) = \gcd(a_n, n) = n$, the last equality being a consequence of the fact that $n \mid a_n$. On the other hand, a_{a_n} is a multiple of a_n , thus $\gcd(a_{a_n}, a_n) = a_n$ and so we obtain $a_n = n$, a contradiction. Thus $a_n = n$ for all n . \square

10. (Iran 2011) Prove that there are infinitely many positive integers n such that $n^2 + 1$ has no proper divisor of the form $k^2 + 1$.

Proof. We say that n is good if $n^2 + 1$ has no proper divisor of the form $k^2 + 1$. We will prove that $F_n = 2^{2^n} + 1$ has a good divisor for all n . Since $\gcd(F_n, F_m) = 1$ for all $n \neq m$, the result follows.

Now, assume that there is n such that all positive divisors of F_n are bad. In particular F_n is bad, hence it has a proper divisor of the form $k^2 + 1$.

This new divisor is also bad, so it has a proper divisor of the form $l^2 + 1$. Continuing like this, we create an infinite decreasing sequence of divisors of F_n , which is clearly absurd. The result follows. \square

11. a) (Romanian Masters in Mathematics 2009) Let a_1, \dots, a_k be nonnegative integers and let $d = \gcd(a_1, \dots, a_k)$ and $n = a_1 + \dots + a_k$. Prove that

$$\frac{d}{n} \cdot \frac{n!}{a_1! \dots a_k!} \in \mathbf{Z}.$$

- b) Prove that $(n)!^k k! \mid (nk)!$ for all positive integers n, k .

Proof. a) Writing $d = a_1 x_1 + \dots + a_k x_k$ for some integer x_1, \dots, x_k , we have

$$\frac{d}{n} \cdot \frac{n!}{a_1! \dots a_k!} = \sum_{i=1}^k x_i \cdot \frac{a_i}{n} \cdot \frac{n!}{a_1! \dots a_k!},$$

thus it suffices to prove that $\frac{a_i}{n} \cdot \frac{n!}{a_1! \dots a_k!}$ is an integer. This is clear if $a_i = 0$, and if $a_i > 0$ we have

$$\frac{a_i}{n} \cdot \frac{n!}{a_1! \dots a_k!} = \frac{(n-1)!}{a_1! \dots a_{i-1}! (a_i-1)! a_{i+1}! \dots a_k!} \in \mathbf{Z},$$

since $b_1! \dots b_k! \mid (b_1 + \dots + b_k)!$ for all nonnegative integers b_1, \dots, b_k (this follows by an immediate induction from the case $k = 2$, which is equivalent to $\binom{b_1+b_2}{b_1} \in \mathbf{Z}$).

- b) We have

$$\frac{(nk)!}{k!(n!)^k} = \prod_{\ell=0}^{k-2} \frac{(n(k-\ell))!}{(k-\ell)n!(n(k-\ell-1))!}$$

and by a) each of the numbers $\frac{(n(k-\ell))!}{(k-\ell)n!(n(k-\ell-1))!}$ is an integer.

One can also give a combinatorial proof, observing that $\frac{(nk)!}{(n!)^k k!}$ equals the number of ways one can divide nk people in k (unordered) groups of n people. \square

12. (Brazil 2011) Are there 2011 positive integers $a_1 < a_2 < \dots < a_{2011}$ such that $\gcd(a_i, a_j) = a_j - a_i$ for any i, j such that $1 \leq i < j \leq 2011$?

Proof. For all $i < j$ we must have $\gcd(a_i, a_j) \leq a_j - a_i$, since $a_j - a_i$ is a positive multiple of $\gcd(a_i, a_j)$. The condition $\gcd(a_i, a_j) = a_j - a_i$ is equivalent to $a_j - a_i \mid a_i$ (as this automatically implies $a_j - a_i \mid a_j$ and so $a_j - a_i \mid \gcd(a_i, a_j)$). We will now prove by induction that for any $n \geq 2$ we can find positive integers $a_1 < \dots < a_n$ such that $a_j - a_i \mid a_i$ for all $i < j$. For $n = 2$ choose $a_1 = 1$ and $a_2 = 2$. Assuming that we have already constructed a_1, \dots, a_n , define $b_1 = a_1 \dots a_n$ and $b_i = a_1 \dots a_n + a_{i-1}$ for $2 \leq i \leq n+1$. Then clearly $b_1 < \dots < b_{n+1}$ and it is not difficult to check that they satisfy $b_j - b_i \mid b_i$ for $i < j$. Indeed, if $i > 1$ this comes down to $a_{j-1} - a_{i-1} \mid a_{i-1} + a_1 \dots a_n$, which is clear since $a_{j-1} - a_{i-1}$ divides both a_{i-1} and $a_1 \dots a_n$. If $i = 1$, this reduces to $a_{j-1} \mid a_1 \dots a_n$, and it is also clear. \square

13. (Tournament of the Towns 2001) Are there positive integers $a_1 < a_2 < \dots < a_{100}$ such that

$$\gcd(a_1, a_2) > \gcd(a_2, a_3) > \dots > \gcd(a_{99}, a_{100}) > \gcd(a_{100}, a_1)?$$

Proof. First, we will build a sequence b_k for $1 \leq k \leq 100$ such that

$$\gcd(b_1, b_2) > \gcd(b_2, b_3) > \dots > \gcd(b_{99}, b_{100}) > \gcd(b_{100}, b_1),$$

not worrying about the relative sizes of the b_k . This is easy. For example, we can take $b_k = (203 - 2k)(205 - 2k)$. Then we compute

$$\begin{aligned} \gcd(b_k, b_{k+1}) &= \gcd((203 - 2k)(205 - 2k), (201 - 2k)(203 - 2k)) \\ &= (203 - 2k) \gcd(205 - 2k, 201 - 2k) \\ &= (203 - 2k) \gcd(205 - 2k, 4) \\ &= 203 - 2k, \end{aligned}$$

since $205 - 2k$ is odd and hence relatively prime to 4. Also

$$\gcd(b_{100}, b_1) = \gcd(15, 201 \cdot 203) = 3 \gcd(5, 67 \cdot 203) = 3,$$

since $67 \cdot 203 = 13601 = 2720 \cdot 5 + 1$ is relatively prime to 5.

Next we fix the relative sizes without changing any of the greatest common divisors. To do this we inductively define $a_1 = b_1$, then $a_k = b_k(1 + a_{k-1}b_{k+1})$ for $2 \leq k \leq 99$, and finally $a_{100} = b_{100}(1 + a_{99}b_1)$. Note that this clearly gives $a_1 < a_2 < \dots < a_{100}$. To see that it doesn't change any of the greatest common divisors, we compute

$$\begin{aligned}\gcd(a_k, a_{k+1}) &= \gcd(a_k, b_{k+1}(1 + a_k b_{k+2})) = \gcd(a_k, b_{k+1}) \\ &= \gcd(b_k(1 + a_{k-1}b_{k+1}), b_{k+1}) = \gcd(b_k, b_{k+1}),\end{aligned}$$

for $2 \leq k \leq 99$. The computation for the remaining cases is similar. \square

14. (Russian Olympiad 2012) Let n be an integer greater than 1. When a runs over all integers greater than 1, what is the maximum number of pairwise relatively prime numbers among $1 + a, 1 + a^2, \dots, 1 + a^{2^n - 1}$?

Proof. We first prove that no more than n of these numbers can be pairwise relatively prime. To do this note that if k is odd then $1 + a^m$ divides $1 + a^{km}$. Note also that each of the numbers $1, 2, 3, \dots, 2^n - 1$ has the form $2^t k$, where $0 \leq t \leq n - 1$ and k is odd. Hence each of the given numbers is divisible by one of the numbers $1 + a, 1 + a^2, 1 + a^4, \dots, 1 + a^{2^{n-1}}$. Therefore among any $n + 1$ of the given numbers there are two which are not relatively prime. Since Fermat numbers are pairwise relatively prime, for $a = 2$ we obtain n pairwise relatively prime numbers, namely $1 + 2, 1 + 2^2, 1 + 2^4, \dots, 1 + 2^{2^{n-1}}$. Hence the desired number is n . \square

15. (Brazilian Olympic Revenge 2014) a) Prove that for all positive integers n we have

$$\gcd\left(n, \left\lfloor n\sqrt{2} \right\rfloor\right) < \sqrt[4]{8n^2}.$$

b) Prove that there are infinitely many positive integers n such that

$$\gcd\left(n, \left\lfloor n\sqrt{2} \right\rfloor\right) > \sqrt[4]{7.99n^2}.$$

Proof. a) Let $d = \gcd\left(n, \lfloor n\sqrt{2} \rfloor\right)$ and write $n = kd$ and $\lfloor n\sqrt{2} \rfloor = md$ for some positive integers k, m . Then

$$md \leq kd\sqrt{2} < md + 1.$$

The first inequality gives $m \leq k\sqrt{2}$ and so $m^2 \leq 2k^2$. This cannot be an equality since $\sqrt{2}$ is irrational, hence $m^2 \leq 2k^2 - 1$. Now the second inequality can be written as

$$d(k\sqrt{2} - m) < 1 \quad \text{or equivalently} \quad d(2k^2 - m^2) < m + k\sqrt{2}.$$

Since $2k^2 - m^2 \geq 1$ and $m \leq k\sqrt{2}$, we obtain $d < 2\sqrt{2}k$, which is equivalent to $d < \sqrt[4]{8n^2}$.

b) Part a) suggests how to take n : with the previous notations, we need to ensure that $2k^2 - m^2 = 1$. This equation has infinitely many solutions in positive integers: the number $(1 + \sqrt{2})^{2N+1}$ can be written as $m_N + k_N\sqrt{2}$, and we have $m_N^2 - 2k_N^2 = -1$. If (m, k) is such a solution, we look for $n = kd$ such that $\lfloor n\sqrt{2} \rfloor = md$, which by the inequalities in part a) is equivalent to $d < k\sqrt{2} + m$. On the other hand, the inequality $\gcd\left(n, \lfloor n\sqrt{2} \rfloor\right) > \sqrt[4]{7.99n^2}$ is equivalent to $d > \sqrt{7.99}k$. But if k is large enough, then we can find an integer d between $\sqrt{7.99}k$ and $k\sqrt{2} + m = k\sqrt{2} + \sqrt{2k^2 - 1}$, and setting $n = kd$ gives a solution of the problem for any such k . \square

16. (AMM) The greatest common divisor of a set D of positive integers is 1. Prove the existence of a bijection $f : \mathbf{Z} \rightarrow \mathbf{Z}$ such that $|f(n) - f(n-1)| \in D$ for all integers n .

Proof. First of all, we claim that we may assume that D is finite. Indeed, if D is infinite, arrange its elements in increasing order $a_1 < a_2 < \dots$.

Setting $x_n = \gcd(a_1, \dots, a_n)$ we have $x_n \geq x_{n+1}$, thus the sequence $(x_n)_{n \geq 1}$ is eventually constant and this constant divides all elements of D , so it must be 1. In other words, D contains a finite subset whose gcd is 1.

Assuming that D is finite, we will prove by induction on the number $|D|$ of elements of D that we can find a bijection $f : \mathbf{Z} \rightarrow \gcd(D) \cdot \mathbf{Z}$ such that $|f(n) - f(n-1)| \in D$ for all n . The case $|D| = 1$ is obvious: if $D = \{d\}$, simply set $f(n) = nd$. Assume now that the result holds for all finite sets of cardinality smaller than k and consider D of cardinality k . Fix some element $b \in D$ and consider $D' = D \setminus \{b\}$. To simplify notations, write $d = \gcd(D)$, $d' = \gcd(D')$ and $k = \frac{d'}{d}$. Applying the inductive hypothesis to D' , we find a bijection $g : \mathbf{Z} \rightarrow d'\mathbf{Z}$ such that $|g(n) - g(n-1)| \in D'$ for all integers n . We will construct the function f in the next paragraph.

Pick any integer n and write $n = qk + r$ with $0 \leq r < k$. Define $f(n) = g(q) + br$ if q is even and $f(n) = g(q) + b(k-1-r)$ if q is odd. It is not difficult to check that any multiple of d can be uniquely written $d'u + br$ with $u \in \mathbf{Z}$ and $0 \leq r < k$ (it suffices to use the equality $d = \gcd(d', b)$). From this it follows immediately that f is bijective. On the other hand, let us check that $|f(n) - f(n-1)| \in D$ for all n . If k does not divide n , then by construction $|f(n) - f(n-1)| = b \in D$. On the other hand, if $n = kx$ is a multiple of k , then again by construction $|f(n) - f(n-1)| = |g(x) - g(x-1)| \in D' \subset D$. This shows that f has all desired properties and finishes the proof. \square

17. (China TST 2012) Let n be an integer greater than 1. Prove that there are only finitely many n -tuples of positive integers (a_1, a_2, \dots, a_n) such that

- a) $a_1 > a_2 > \dots > a_n$ and $\gcd(a_1, a_2, \dots, a_n) = 1$;
 b) $a_1 = \gcd(a_1, a_2) + \gcd(a_2, a_3) + \dots + \gcd(a_{n-1}, a_n) + \gcd(a_n, a_1)$.

Proof. The essential part consists of course in understanding what condition b) really says. Since $\gcd(a_i, a_{i+1}) \leq a_i - a_{i+1}$ for $1 \leq i < n$ (this uses part a)), it follows that

$$\gcd(a_1, a_2) + \dots + \gcd(a_n, a_1) \leq a_1 - a_2 + a_2 - a_3 + \dots + a_{n-1} - a_n + \gcd(a_n, a_1)$$

hence $a_1 \leq a_1 - a_n + \gcd(a_n, a_1)$ and $a_n \leq \gcd(a_n, a_1)$. Since $a_n \geq \gcd(a_n, a_1)$, this means that all previous inequalities must be equalities. Thus $a_i = a_{i+1} + \gcd(a_i, a_{i+1})$ for $1 \leq i < n$ and $a_n \mid a_1$.

Let $b_i = \frac{a_{i+1}}{\gcd(a_i, a_{i+1})}$ for $1 \leq i < n$. Then $a_i = a_{i+1}(1 + \frac{1}{b_i})$ and so

$$\frac{a_1}{a_n} = \left(1 + \frac{1}{b_1}\right) \cdots \left(1 + \frac{1}{b_{n-1}}\right)$$

is an integer. Note that this integer is less than or equal to 2^{n-1} since each factor of the product is less than or equal to 2. The following lemma implies that there are only finitely many such tuples (b_1, \dots, b_{n-1}) .

Lemma 8.5. *For any positive real number x and any positive integer k there are only finitely many (maybe zero) k -tuples (b_1, \dots, b_k) of positive integers such that*

$$\left(1 + \frac{1}{b_1}\right) \cdot \left(1 + \frac{1}{b_2}\right) \cdots \left(1 + \frac{1}{b_k}\right) = x.$$

Proof. This is easily proved by induction on k , the assertion being clear for $k = 1$. Assume that it holds for $k - 1$ and let us prove it for k . Of course, we may assume that $x > 1$, as otherwise there is no solution. If

$$\left(1 + \frac{1}{b_1}\right) \cdot \left(1 + \frac{1}{b_2}\right) \cdots \left(1 + \frac{1}{b_k}\right) = x,$$

then some b_i must satisfy $1 + \frac{1}{b_i} > \sqrt[k]{x}$, and this b_i can only take finitely many possible values. By the inductive hypothesis, for each possible value of b_i we can find only finitely many $k - 1$ -tuples $(b_j)_{j \neq i}$ satisfying

$$\prod_{j \neq i} \left(1 + \frac{1}{b_j}\right) = \frac{x}{1 + \frac{1}{b_i}},$$

yielding the desired result. □

We haven't used so far the hypothesis $\gcd(a_1, \dots, a_n) = 1$. Note that $b_1 b_2 \dots b_{n-1} a_i$ is a multiple of a_n for all $1 \leq i \leq n$. Thus a_n must divide

$$b_1 b_2 \dots b_{n-1} \gcd(a_1, \dots, a_n) = b_1 b_2 \dots b_{n-1}$$

and so a_n can only take finitely many values. Since b_1, b_2, \dots, b_{n-1} also take only finitely many values, it follows that all a_i 's have the same property and the problem is solved. \square

18. Integers a, b and rational numbers x, y satisfy $y^2 = x^3 + ax + b$. Prove that we can write $x = \frac{u}{v^2}$ and $y = \frac{w}{v^3}$ for some integers u, v, w , with $\gcd(u, v) = \gcd(w, v) = 1$.

Proof. Write $x = \frac{p}{q}$ and $y = \frac{r}{s}$ with p, q, r, s integers, $q, s > 0$ and $\gcd(p, q) = \gcd(r, s) = 1$. Clearing denominators, the equation

$$y^2 = x^3 + ax + b$$

is equivalent to

$$r^2 q^3 = p^3 s^2 + apq^2 s^2 + bq^3 s^2.$$

The right-hand side is a multiple of s^2 , hence $s^2 \mid r^2 q^3$. Since $\gcd(r, s) = 1$, it follows that $s^2 \mid q^3$. On the other hand, taking the equation modulo q^3 , we obtain $ps^2(p^2 + aq^2) \equiv 0 \pmod{q^3}$. Since $\gcd(q, p) = 1$, we have $\gcd(q, p(p^2 + aq^2)) = 1$, hence the previous congruence yields $q^3 \mid s^2$. We conclude that $q^3 = s^2$. But then q is a square, say $q = v^2$, and necessarily $s = v^3$. The result follows. \square

19. (Kvant M 905) Let x and n be positive integers such that $4x^n + (x+1)^2$ is a perfect square. Prove that $n = 2$ and find at least one x with this property.

Proof. Let $4x^n + (x+1)^2 = y^2$. Then

$$(y - x - 1)(y + x + 1) = 4x^n$$

and since $y - x - 1$ and $y + x + 1$ have equal parity we conclude they are even. Set $y - x - 1 = 2a$. Then $y + x + 1 = 2(a + x + 1)$ and we get $a(a + x + 1) = x^n$. But a and $a + x + 1$ are relatively prime since otherwise $x + 1$ and x^n would have a common divisor. Hence $a = u^n$, $a + x + 1 = v^n$, $x = uv$ and therefore $uv + 1 = v^n - u^n$. But this is not possible for $n = 1$ (since $vu > v - u$) or $n \geq 3$ since in this case

$$v^n - u^n = (v - u)(v^{n-1} + v^{n-2}u + \cdots + u^{n-1}) \geq uv + 2.$$

Hence $n = 2$ and for $x = 2$ we have that $y = 5$ (this is not the only solution, for instance for $x = 104$ we have that $y = 233$). \square

20. Solve in positive integers the equation

$$\frac{1}{x^2} + \frac{1}{y^2} = \frac{1}{z^2}.$$

Proof. The equation is equivalent to

$$x^2 + y^2 = \left(\frac{xy}{z}\right)^2.$$

Since $z^2 \mid (xy)^2$, we have $z \mid xy$, hence there is a positive integer t such that $xy = zt$, and then the previous equation yields $x^2 + y^2 = t^2$. By theorem 3.50 and by symmetry in x and y we may write

$$x = d(m^2 - n^2), \quad y = 2dmn, \quad t = d(m^2 + n^2)$$

with $m > n > 0$ of different parities and relatively prime. Then $xy = zt$ can be written as

$$z(m^2 + n^2) = 2dmn(m^2 - n^2).$$

Note that $m^2 + n^2$ is odd and relatively prime to $m, n, m^2 - n^2$, since m, n are relatively prime and of different parities. Thus $m^2 + n^2$ must divide d . Writing $d = k(m^2 + n^2)$ and recalling that $z = \frac{xy}{t}$ we obtain the solutions

$$x = k(m^4 - n^4), \quad y = 2kmn(m^2 + n^2), \quad z = 2kmn(m^2 - n^2)$$

and (recalling the symmetry in x and y)

$$x = 2kmn(m^2 + n^2), \quad y = k(m^4 - n^4), \quad z = 2kmn(m^2 - n^2). \quad \square$$

21. (Romania TST 2015) A Pythagorean triple is a solution (x, y, z) of the equation $x^2 + y^2 = z^2$ in positive integers, where we count (x, y, z) and (y, x, z) as the same triple. Given a non-negative integer n , prove that some positive integer appears in precisely n distinct Pythagorean triples.

Proof. We will prove that 3^n appears in precisely n Pythagorean triples. This is clear when $n = 0$, so assume that $n > 0$. First, the equation $x^2 + y^2 = 3^{2n}$ does not have solutions with $x, y > 0$. Indeed, it is not difficult to see that x, y must be multiples of 3, thus $x = 3x_1, y = 3y_1$ and $x_1^2 + y_1^2 = 3^{2(n-1)}$, thus we can repeat the argument and obtain positive integers x_n, y_n such that $x_n^2 + y_n^2 = 1$, which is obviously impossible.

Let us deal now with the equation $3^{2n} + y^2 = z^2$. Then $(z - y)(z + y) = 3^{2n}$, thus $z - y = 3^a$ and $z + y = 3^b$ with $a + b = 2n$. This gives us $y = \frac{3^b - 3^a}{2}$ and $z = \frac{3^b + 3^a}{2}$. Note that since $y > 0$, we must have $b > a$. Conversely, for each $b \in \{n + 1, \dots, 2n\}$ setting $a = 2n - b$ and defining y, z by the formulae above we obtain a solution. We obtain thus exactly n Pythagorean triples containing 3^n . \square

22. Find all triples (x, y, n) of positive integers with $\gcd(x, n + 1) = 1$ and $x^n + 1 = y^{n+1}$.

Proof. If $n = 1$ we obtain $x = y^2 - 1$ and since x must be odd, y can be any even positive number. Assume that $n > 1$ and that $x^n + 1 = y^{n+1}$, with $\gcd(x, n + 1) = 1$. Then

$$(y - 1)(y^n + y^{n-1} + \dots + y + 1) = x^n.$$

If d is a common divisor of $y - 1$ and $y^n + \dots + y + 1$, then d divides $n + 1$ (since $y^n + \dots + y + 1 \equiv n + 1 \pmod{y - 1}$) and d divides x^n , but then $d \mid \gcd(x^n, n + 1) = 1$. Thus $y - 1$ and $y^n + \dots + y + 1$ are relatively prime. Since their product is an n th power, we deduce that both are

n th powers. Say $y^n + \dots + y + 1 = a^n$ for some positive integer a . Since $n > 1$, the binomial formula shows that

$$y^n < y^n + \dots + y + 1 < (y + 1)^n,$$

yielding $y < a < y + 1$, a contradiction. Thus there are no solutions with $n \geq 2$. \square

23. Let n be a positive integer such that n^2 is the difference of the cubes of two consecutive positive integers. Prove that n is the sum of the squares of two consecutive positive integers.

Proof. Let $n^2 = (m + 1)^3 - m^3$. Then n is odd and $n^2 = 3m^2 + 3m + 1$ which can be written as $(2n + 1)(2n - 1) = 3(2m + 1)^2$. Since $2n - 1$ and $2n + 1$ are relatively prime it follows that one of them is a perfect square. But n is odd and $2n + 1 \equiv 3 \pmod{4}$, so $2n + 1$ is not a perfect square. Hence $2n - 1 = (2l + 1)^2$ and $n = l^2 + (l + 1)^2$. \square

24. (Vietnam 2007) Let x, y be integers different from -1 such that $\frac{x^4-1}{y+1} + \frac{y^4-1}{x+1}$ is also an integer. Prove that $x^4 y^{44} - 1$ is a multiple of $x + 1$.

Proof. Let $a = \frac{x^4-1}{y+1}$ and $b = \frac{y^4-1}{x+1}$. By assumption a, b are rational numbers and $a + b$ is an integer. Note that $ab = \frac{x^4-1}{x+1} \cdot \frac{y^4-1}{y+1}$ is also an integer, since $u^4 - 1$ is a multiple of $u + 1$ for any integer u . Thus the polynomial $(X - a)(X - b) = X^2 - (a + b)X + ab$ has integer coefficients and rational roots a, b . We deduce that a, b are integers, thus $x + 1 \mid y^4 - 1$. Then clearly $x + 1 \mid y^{44} - 1$ and since $x^4 \equiv 1 \pmod{x + 1}$, the result follows. \square

25. (Balkan 2006) Find all triplets of positive rational numbers (m, n, p) such that the numbers $m + \frac{1}{np}$, $n + \frac{1}{pm}$, $p + \frac{1}{mn}$ are all integers.

Proof. Clearly mnp plays a key role in the problem, so denote $a = mnp$. By assumption $\frac{a+1}{np}$, $\frac{a+1}{pm}$, $\frac{a+1}{mn}$ are integers, hence so is their product, i.e.

$\frac{(a+1)^3}{a^2}$ is an integer. Write $(a+1)^3 = ka^2$ for some integer k , then a is a rational root of the monic polynomial with integer coefficients $(X+1)^3 - kX^2$. Thus a is an integer. But then $a \mid ka^2 = (a+1)^3$, thus $a \mid 1$ and so $a = 1$. It follows that $\frac{a+1}{np} = \frac{2}{np} = 2m$ is an integer and similarly $2n$ and $2p$ are integers. Moreover, the product of $2m, 2n, 2p$ equals 8. Considering the possible decompositions of 8 as a product of three positive integers, we obtain the solutions $(1, 1, 1), (4, \frac{1}{2}, \frac{1}{2}), (2, \frac{1}{2}, 1)$ and their permutations. \square

26. A polynomial f has integer coefficients and satisfies $|f(a)| = |f(b)| = 1$ for some distinct integers a, b .

a) Prove that if $|a - b| > 2$, then f has no rational root.

b) Prove that if $|a - b| = 2$, then the only possible rational root of f is $\frac{a+b}{2}$.

Proof. a) Assume that $x = \frac{p}{q}$ is a rational root of f , with p, q relatively prime integers. By example 3.64 we know that we can write

$$f(X) = (qX - p)g(X)$$

for some polynomial g with integer coefficients. Then

$$|(qa - p)| \cdot |g(a)| = |f(a)| = 1$$

and so $|qa - p| = |g(a)| = 1$. Similarly $|qb - p| = 1$. But then

$$|qa - qb| = |(qa - p) - (qb - p)| \leq |qa - p| + |qb - p| = 2,$$

thus $|a - b| \leq 2$ (as $|q| \geq 1$), a contradiction.

b) We still obtain $|qa - qb| \leq 2$ and since $|a - b| = 2$ we must have $|q| \leq 1$. Since trivially $|q| \geq 1$, we deduce that all previous inequalities must be equalities. In particular $|q| = 1$ and the numbers $qa - p$ and $p - qb$ must have the same sign. Since both have absolute value 1, we must have $qa - p = p - qb$, thus $x = \frac{p}{q} = \frac{a+b}{2}$, as desired. \square

27. (Turkey 2003) Find all positive integers n for which $2^{2n+1} + 2^n + 1$ is a perfect power.

Proof. Assume that $2^{2n+1} + 2^n + 1 = a^k$ for some integers $a, k > 1$. First, assume that k is even and let $b = a^{\frac{k}{2}}$, so that

$$2^n(2^{n+1} + 1) = b^2 - 1 = (b - 1)(b + 1).$$

Since $\gcd(b - 1, b + 1) = 2$, we deduce that $2^{n-1} \mid b - 1$ or $2^{n-1} \mid b + 1$. Write $b - r = 2^{n-1}c$ with $r \in \{-1, 1\}$ and $c > 0$. The previous equality is equivalent (after division by 2^n) to

$$2^{n+1} + 1 = c(r + 2^{n-2}c),$$

or equivalently

$$2^{n-2}(c^2 - 8) + cr - 1 = 0.$$

Hence $c^2 - 8 \mid cr - 1 \mid c^2 - 1$ and so $c^2 - 8 \mid 7$. This easily implies $c = 3$ (the case $c = 1$ is easily excluded by going back to the equation) and then $r = -1$ and $n = 4$, which is indeed a solution of the problem, and the only solution for which k is even.

Suppose now that k is odd. Then

$$2^n(2^{n+1} + 1) = a^k - 1 = (a - 1)(1 + a + \dots + a^{k-1}).$$

Clearly a is odd, hence $1 + a + \dots + a^{k-1}$ is also odd. The previous relation implies that $2^n \mid a - 1$ and $1 + a + \dots + a^{k-1} \mid 2^{n+1} + 1$. Thus $a \geq 2^n + 1$ and $1 + a + \dots + a^{k-1} \leq 2^{n+1} + 1$. But

$$1 + a + \dots + a^{k-1} \geq 1 + a + a^2 > 1 + 2^n + 2^{2n} > 1 + 2^{n+1},$$

a contradiction. Hence there are no solutions in this case and $n = 4$ is the only solution of the problem. \square

Remark 8.6. The equation $2^{2n+1} + 2^n + 1 = x^2$ was proposed at IMO 2006 (!).

28. Let f be a polynomial with rational coefficients such that for all positive integers n the equation $f(x) = n$ has at least one rational solution. Prove that $\deg(f) = 1$.

Proof. Clearly f cannot be constant, so assume that $d = \deg(f) > 1$. Let x_n be a rational solution of the equation $f(x_n) = n$. Choose a positive integer N such that the polynomial $Nf = g$ has integer coefficients. Then $g(x_n) = nN$ and by the rational root theorem the denominator of x_n (when written in lowest form) divides the leading coefficient C of g . Letting $a_n = Cx_n$, we obtain a sequence of integers a_n such that $g\left(\frac{a_n}{C}\right) = nN$. Note that $a_n \neq a_m$ for all $n \neq m$ by the previous equality. Thus a_1, \dots, a_n are pairwise distinct integers, and so any positive integer appears at most twice among $|a_1|, |a_2|, \dots$. On the other hand, since $\deg(g) = d > 1$, there is M such that for $|x| > M$ we have $|g(x)| \geq x^2$. For n large enough we have $|a_n| > cM$ and so

$$nN = \left| g\left(\frac{a_n}{C}\right) \right| \geq \left| \frac{a_n}{C} \right|^2.$$

We deduce that $|a_n| \leq c\sqrt{nN} = D\sqrt{n}$. But then among $|a_1|, |a_2|, \dots, |a_n|$ there can be at most $D\sqrt{n} + E$ (E being another constant independent of n) distinct integers, contradicting the fact that each positive integer appears at most twice in this sequence. Thus $d = 1$. \square

29. (Kyiv mathematical festival 2014)

a) Let y be a positive integer. Prove that for infinitely many positive integers x we have

$$\text{lcm}(x, y+1) \cdot \text{lcm}(x+1, y) = x(x+1).$$

b) Prove that there exists positive integer y such that

$$\text{lcm}(x, y+1) \cdot \text{lcm}(x+1, y) = y(y+1)$$

for at least 2014 positive integers x .

Proof. a) Note that $\text{lcm}(x, y+1)$ is a multiple of x and $\text{lcm}(x+1, y)$ is a multiple of $x+1$, thus the equality in the statement of the problem is equivalent to the simultaneous equalities $\text{lcm}(x, y+1) = x$ and $\text{lcm}(x+1, y) = x+1$, i.e. to $y+1 \mid x$ and $y \mid x+1$. Look for $x = k(y+1)$, the condition $y \mid x+1$ is equivalent to $y \mid ky + k + 1$, or $y \mid k + 1$. It is thus enough to take $x = (ry - 1)(y + 1)$ for $r > 1$.

b) The same remarks as in part a) show that the equality is satisfied if and only if $x \mid y+1$ and $x+1 \mid y$. Taking $y = 2^{2^N} - 1$ with N large enough, any $x = 2^{2^d}$ with $1 \leq d \leq N-1$ satisfies $x \mid y+1$ and $x+1 \mid y$. \square

30. (Kvant M 666) Find the least positive integer a for which there exist pairwise different positive integers a_1, a_2, \dots, a_9 greater than a such that

$$\text{lcm}(a, a_1, a_2, \dots, a_9) = 10a.$$

Proof. We may assume that $a < a_1 < \dots < a_9$.

Set $A = \text{lcm}(a, a_1, a_2, \dots, a_9)$. Then

$$\frac{A}{a} > \frac{A}{a_1} > \dots > \frac{A}{a_9}$$

are positive integers and therefore $\frac{A}{a} \geq 10$. Since $\frac{A}{a} = 10$ we conclude that

$$\frac{A}{a} = 10, \quad \frac{A}{a_1} = 9, \quad \dots, \quad \frac{A}{a_9} = 1.$$

Hence

$$a_9 = A, \quad a_8 = \frac{A}{2}, \dots, \quad a_1 = \frac{A}{9}, \quad a = \frac{A}{10}$$

and A is divisible by $\text{lcm}(2, 3, \dots, 10) = 2^3 3^2 \cdot 5 \cdot 7$. The least a is equal to $\frac{2^3 3^2 \cdot 5 \cdot 7}{10} = 252$. In this case the numbers $a_k = \frac{2^3 3^2 \cdot 5 \cdot 7}{10-k}$, $k = 1, 2, \dots, 9$ satisfy the given condition. \square

31. (Korea 2013) Find all functions $f : \mathbf{N} \rightarrow \mathbf{N}$ satisfying

$$f(mn) = \text{lcm}(m, n) \cdot \gcd(f(m), f(n))$$

for all positive integers m, n .

Proof. Taking $m = 1$ and setting $a = f(1)$ we obtain

$$f(n) = n \cdot \gcd(a, f(n)).$$

In particular $n \mid f(n)$ for all n . Next, replacing n by an we obtain (taking into account that $a \mid f(an)$)

$$f(an) = an \cdot \gcd(a, f(an)) = a^2 n.$$

Finally, replacing n by an in the original relation, we obtain

$$f(amn) = \text{lcm}(m, an) \cdot \gcd(f(m), f(an)),$$

which can be rewritten

$$a^2 mn = \frac{amn}{\gcd(m, an)} \cdot \gcd(f(m), f(an)).$$

Dividing this last relation by amn , we deduce that $a \mid \gcd(f(m), f(an))$ and so $a \mid f(m)$ for all m . But then $\gcd(a, f(n)) = a$ and so $f(n) = n \cdot \gcd(a, f(n)) = an$ for all n . Conversely, it is not difficult to see that for any positive integer a setting $f(n) = an$ we obtain a solution of the problem. \square

32. (Romania TST 1995) Let $f(n) = \text{lcm}(1, 2, \dots, n)$. Prove that for any $n \geq 2$ one can find a positive integer x such that

$$f(x) = f(x+1) = \dots = f(x+n).$$

Proof. It suffices to find x such that $x+1, x+2, \dots, x+n$ are all divisors of $\text{lcm}(1, 2, \dots, x) = f(x)$. Choose $x = 1 + N!$ for some N to be chosen later. Then for all $j \in \{1, 2, \dots, n\}$ we have

$$x + j = j + 1 + N! = (j + 1) \left(\frac{N!}{j + 1} + 1 \right).$$

If we manage to ensure that $j+1$ and $\frac{N!}{j+1} + 1$ are relatively prime integers between 1 and x , it will follow that $x + j$ divides $f(x)$. But this is very easy to realize: simply take N such that $N!$ is a multiple of $(j + 1)^2$ for all $j \leq n$, which is certainly possible. \square

33. Prove that for all positive integers a_1, \dots, a_n

$$\text{lcm}(a_1, \dots, a_n) \geq \frac{a_1 a_2 \dots a_n}{\prod_{1 \leq i < j \leq n} \gcd(a_i, a_j)}.$$

Proof. If $n = 2$, the desired inequality is an equality. Next, we prove the result by induction. Assuming that it holds for $n - 1$, denote $m = \text{lcm}(a_1, \dots, a_{n-1})$ and observe that

$$\text{lcm}(a_1, \dots, a_n) = \text{lcm}(m, a_n) = \frac{m a_n}{\gcd(m, a_n)}.$$

Using the inductive hypothesis, we are reduced to proving that

$$\frac{a_n}{\gcd(m, a_n)} \cdot \frac{a_1 \dots a_{n-1}}{\prod_{1 \leq i < j \leq n-1} \gcd(a_i, a_j)} \geq \frac{a_1 \dots a_n}{\prod_{1 \leq i < j \leq n} \gcd(a_i, a_j)},$$

or equivalently

$$\gcd(m, a_n) \leq \prod_{i=1}^{n-1} \gcd(a_n, a_i).$$

But using exercise 1 (more precisely an $n - 1$ -variable version of it, which follows directly from the cited corollary and an obvious induction) we obtain

$$\gcd(m, a_n) \leq \gcd(a_1 \dots a_{n-1}, a_n) \leq \prod_{i=1}^{n-1} \gcd(a_n, a_i). \quad \square$$

34. (AMM 3834) Let $n > 4$ and let $a_1 < a_2 < \cdots < a_n \leq 2n$ be positive integers. Prove that

$$\min_{1 \leq i \neq j \leq n} \text{lcm}(a_i, a_j) \leq 6(\lfloor n/2 \rfloor + 1).$$

Proof. The key (simple) observation is that for any $1 \leq i \leq n$ we can find a positive integer k_i such that $k_i a_i \in \{n+1, \dots, 2n\}$. Indeed, if $a_i > n$ simply choose $k_i = 1$, otherwise since $\frac{2n}{a_i} - \frac{n}{a_i} \geq 1$ there is an integer k_i between $\frac{n}{a_i}$ and $\frac{2n}{a_i}$.

Using this observation, it is not difficult to conclude: if $k_i a_i = k_j a_j$ for some $i \neq j$, then $k_i a_i$ is a common multiple of a_i and a_j , thus $\text{lcm}(a_i, a_j) \leq k_i a_i \leq 2n$ and we are done (with an even better bound). If this never happens, then the pairwise distinct numbers $k_1 a_1, \dots, k_n a_n$ between $n+1$ and $2n$ must be a permutation of $n+1, \dots, 2n$. Since $n > 4$, we have $3(\lfloor \frac{n}{2} \rfloor + 1) \in \{n+1, \dots, 2n\}$ ($3(\lfloor \frac{n}{2} \rfloor + 1)$ is clearly greater than $\frac{3n}{2} > n$ and smaller than or equal to $\frac{3n}{2} + 3$ and this is smaller than or equal to $2n$ for $n \geq 6$; one easily checks the claim for $n = 5$). Similarly $2(\lfloor \frac{n}{2} \rfloor + 1) \in \{n+1, \dots, 2n\}$, thus there are indices i, j such that $k_i a_i = 2(\lfloor \frac{n}{2} \rfloor + 1)$ and $k_j a_j = 3(\lfloor \frac{n}{2} \rfloor + 1)$. But then $\text{lcm}(a_i, a_j)$ divides $6(\lfloor \frac{n}{2} \rfloor + 1)$ and the result follows. \square

Remark 8.7. The result does not hold for $n = 4$: consider the numbers 5, 6, 7, 8. On the other hand, it is not difficult to check that it holds for $n \leq 3$. The expression $6(\lfloor \frac{n}{2} \rfloor + 1)$ is optimal, since one can check without too much difficulty that we have equality for the sequence $n+1, n+2, \dots, n+n$.

35. Let $(a_n)_{n \geq 1}$ be a sequence of integers such that $m - n \mid a_m - a_n$ for all $m, n \geq 1$. Suppose that there is a polynomial f such that $|a_n| \leq f(n)$ for all $n \geq 1$. Prove that there is a polynomial P with rational coefficients such that $a_n = P(n)$ for all $n \geq 1$.

Proof. Let $d = \deg f$ and define

$$P(X) = \sum_{k=1}^{d+1} a_k \prod_{j \neq k} \frac{X - j}{k - j}.$$

This intimidating polynomial is the unique polynomial of degree $\leq d$ such that $P(n) = a_n$ for $1 \leq n \leq d+1$. We will prove that $a_n = P(n)$ for all n .

Note that P has rational coefficients, so we can find a positive integer N such that all coefficients of NP are integers. Consider the sequence $(b_n)_{n \geq 1}$ defined by $b_n = Na_n - NP(n)$. It is a sequence of integers and it satisfies $m - n \mid b_m - b_n$ for all n (since the sequences $(Na_n)_{n \geq 1}$ and $(NP(n))_{n \geq 1}$ have this property, the first by assumption and the second since NP has integer coefficients.). Since $b_1 = \dots = b_{d+1} = 0$, this implies that $n - 1, \dots, n - (d + 1)$ all divide b_n , thus

$$\text{lcm}(n - 1, \dots, n - d - 1) \mid b_n.$$

On the other hand, exercise 33 yields the existence of a constant $C(d)$ (depending only on d) such that for all $n > d + 1$

$$\text{lcm}(n - 1, \dots, n - d - 1) \geq C(d)n^{d+1}.$$

Since $\deg f, \deg P \leq d$, we have

$$|b_n| \leq Nf(n) + N|P(n)| < C(d)n^{d+1} \leq \text{lcm}(n - 1, \dots, n - d - 1)$$

for n large enough. Thus we must have $b_n = 0$ for n large enough, say for $n \geq M$. But then for any $n \geq 1$ and $m \geq M$ we have $m - n \mid b_m - b_n = b_n$, thus necessarily $b_n = 0$ and so $a_n = P(n)$ for all n . \square

36. Let n, k be positive integers and let $1 < a_1 < \dots < a_k \leq n$ be a sequence of integers such that $\text{lcm}(a_i, a_j) \leq n$ for all $1 \leq i, j \leq k$. Prove that $k \leq 2 \lfloor \sqrt{n} \rfloor$.

Proof. We have

$$n \geq \text{lcm}(a_i, a_{i+1}) = \frac{a_i a_{i+1}}{\gcd(a_i, a_{i+1})} \geq \frac{a_i a_{i+1}}{a_{i+1} - a_i},$$

which can also be written as

$$\frac{1}{a_i} - \frac{1}{a_{i+1}} \geq \frac{1}{n}$$

for $1 \leq i < k$. Given $1 \leq j \leq k-1$, we obtain

$$\sum_{i=j}^{k-1} \left(\frac{1}{a_i} - \frac{1}{a_{i+1}} \right) \geq \frac{k-j}{n},$$

which simplifies to $\frac{1}{a_j} - \frac{1}{a_k} \geq \frac{k-j}{n}$. Since $a_k \leq n$, this last inequality yields $a_j \leq \frac{n}{k-j+1}$. On the other hand, since $a_j > a_{j-1} > \dots > a_1 \geq 1$, we must have $a_j \geq j$. We conclude that for all $1 \leq j < k$ we have $j(k-j+1) \leq n$. Write $k+1 = 2q+r$ for some $r \in \{0, 1\}$ and some $q \geq 1$ (if $q=0$, then $k < 2$ and we are done). Then $q < k$, hence $q(k+1-q) \leq n$. This yields $q^2 \leq n$, hence $q \leq \lfloor \sqrt{n} \rfloor$ and then $k \leq 2q \leq 2 \lfloor \sqrt{n} \rfloor$. \square

37. (AMM E 3350) For $n \geq 1$ and $1 \leq k \leq n$ define

$$A(n, k) = \text{lcm}(n, n-1, \dots, n-k+1).$$

Let $f(n)$ be the largest k such that $A(n, 1) < A(n, 2) < \dots < A(n, k)$.

a) Prove that $f(n) \leq 3\sqrt{n}$.

b) Prove that $f(n) > k$ if $n > k! + k$.

Proof. We need to make a few observations before embarking on the proof. The first and most important observation is that since

$$A(n, k+1) = \text{lcm}(n-k, A(n, k)),$$

we always have $A(n, k+1) \geq A(n, k)$, with equality if and only if $n-k$ divides $A(n, k)$. We deduce that if $A(n, k) = A(n, k+1)$, then

$$A(n+j, k+j) = A(n+j, k+j+1)$$

for all $j \geq 1$ and so $f(n+j) \leq f(n) + j$ for all $n, j \geq 1$.

a) We claim that it suffices to prove that $f(n^2) \leq n$ for all n . Indeed, if this happens, then for any n we can find k such that $k^2 \leq n < (k+1)^2$, thus

$$f(n) \leq f(k^2) + n - k^2 \leq k + n - k^2 \leq k + k^2 + 2k - k^2 = 3k \leq 3\sqrt{n},$$

as needed. In order to prove that $f(n^2) \leq n$, it suffices to prove that $A(n^2, n) = A(n^2, n+1)$, or equivalently that $n^2 - n \mid A(n^2, n)$. This is very simple, since $n^2 - n$ already divides $A(n^2, 2) = n^2(n^2 - 1)$.

b) We have

$$\begin{aligned} A(n, k+1) &= \text{lcm}(n-k, A(n, k)) = \frac{(n-k)A(n, k)}{\gcd(n-k, A(n, k))} \\ &\geq \frac{(n-k)A(n, k)}{\gcd(n-k, n) \cdot \gcd(n-k, n-1) \cdots \gcd(n-k, n-k+1)} \\ &\geq \frac{A(n, k) \cdot (n-k)}{k!}. \end{aligned}$$

Thus for $n > k! + k$ we have $A(n, 1) < \dots < A(n, k+1)$ and so $f(n) > k$. \square

38. Let $a_1 < a_2 < \dots < a_n$ be an arithmetic progression of positive integers such that a_1 is relatively prime to the common difference. Prove that $a_1 a_2 \dots a_n$ divides $(n-1)! \cdot \text{lcm}(a_1, \dots, a_n)$.

Proof. Let d be the common difference, so $a_i = a_1 + (i-1)d$ for $1 \leq i \leq n$. The key ingredient is the identity

$$\frac{d^{n-1}(n-1)!}{a_1 a_2 \dots a_n} = \sum_{k=1}^n (-1)^{k-1} \frac{\binom{n-1}{k-1}}{a_1 + (k-1)d}.$$

This follows from the identity

$$\frac{n!}{(x+1)\dots(x+n)} = \sum_{k=1}^n (-1)^{k-1} \frac{k \binom{n}{k}}{x+k},$$

that has already been established during the solution of practice problem 36 in chapter 1, by letting $x = \frac{a_1}{d}$ and by observing that $k \binom{n}{k} = n \binom{n-1}{k-1}$. The right-hand side is clearly of the form $\frac{s}{\text{lcm}(a_1, \dots, a_n)}$ for some integer s . Thus $a_1 a_2 \dots a_n$ divides $d^{n-1}(n-1)! \text{lcm}(a_1, \dots, a_n)$. But a_1 and hence all the a_i are relative prime to d , so we may cancel off the factor of d^{n-1} . The result follows. \square

39. Let $n > 1$ and let $a_0 < a_1 < \dots < a_n$ be positive integers such that $\frac{1}{a_0}, \dots, \frac{1}{a_n}$ is an arithmetic progression. Prove that

$$a_0 \geq \frac{2^n}{n+1}.$$

Proof. Let $M = \text{lcm}(a_0, \dots, a_n)$ and write $M = a_i b_i$ for positive integers $b_0 > b_1 > \dots > b_n$. By assumption b_0, \dots, b_n form an arithmetic progression and $b_i \mid M$ for all i , thus $M \geq \text{lcm}(b_0, \dots, b_n)$ and so

$$a_0 \geq \frac{\text{lcm}(b_0, \dots, b_n)}{b_0}.$$

It suffices to prove that for any arithmetic progression $b_n < \dots < b_0$ of positive integers we have

$$\frac{\text{lcm}(b_0, \dots, b_n)}{b_0} \geq \frac{2^n}{n+1}.$$

Let d be the common difference of $b_n < \dots < b_0$. Dividing each b_i by $\gcd(d, b_n)$ does not change the quotient $\frac{\text{lcm}(b_0, \dots, b_n)}{b_0}$, thus we may assume that $\gcd(d, b_n) = 1$, in which case $\gcd(d, b_i) = 1$ for all i (since $b_i = b_n + (n-i)d$). Thus $\gcd(d, b_0 \dots b_n) = 1$ and so $\gcd(d, b_0 \dots b_k) = 1$ for all $k \leq n$. Let $k = \lfloor \frac{n}{2} \rfloor$ and apply the previous exercise to the arithmetic progression $b_0 > b_1 > \dots > b_k$. Since $\gcd(d, b_0 \dots b_k) = 1$, we deduce that

$$\frac{\text{lcm}(b_0, \dots, b_n)}{b_0} \geq \frac{\text{lcm}(b_0, \dots, b_k)}{b_0} \geq \frac{b_1 \dots b_k}{k!}.$$

Observe now that $b_n \geq 1$, $b_{n-1} \geq 2$, in general $b_j \geq n - j + 1$, thus $b_k \geq n - k + 1, \dots, b_1 \geq n$. Thus

$$\frac{b_1 \dots b_k}{k!} \geq \frac{n(n-1) \dots (n-k+1)}{k!} = \binom{n}{k}.$$

Since the binomial coefficient $\binom{n}{k}$ is the largest among $\binom{n}{i}$ with $0 \leq i \leq n$, and since these binomial coefficients add up to 2^n , we have

$$\binom{n}{k} \geq \frac{2^n}{n+1}.$$

The result follows. □

8.3 The fundamental theorem of arithmetic

1. Prove that if a is an integer greater than 1 and if $n > 1$ is not a power of 2, then $a^n + 1$ is composite.

Proof. Since n is not a power of 2, we can write $n = 2^k \cdot m$ with $m > 1$ odd and $k \geq 0$. Then $a^{2^k} + 1$ divides $(a^{2^k})^m + 1 = a^n + 1$ and $1 < a^{2^k} + 1 < a^n + 1$, hence $a^n + 1$ is composite. \square

2. (St. Petersburg 2004) Prove that for any integer a there exist infinitely many positive integers n such that $a^{2^n} + 2^n$ is composite.

Proof. If $a = 0$ we can choose any integer $n > 1$, so assume that $a \neq 0$. Replacing a with $-a$, we may assume that $a > 0$. If $a = 1$ choose any $n > 1$ which is not a power of 2 and use the previous exercise, so assume that $a > 1$. Then choose any odd integer $k > 1$ and set $n = 2k$. We have

$$a^{2^n} + 2^n = a^{2^n} + 4 \cdot 4^{k-1} = x^4 + 4y^4,$$

where $x = a^{2^{n-2}}$ and $y = 2^{\frac{k-1}{2}}$. Note that $x, y > 1$ and

$$x^4 + 4y^4 = (x^2 + 2y^2)^2 - (2xy)^2 = (x^2 - 2xy + 2y^2)(x^2 + 2xy + 2y^2)$$

is composite. \square

3. Find all positive integers n for which at least one of the numbers $n^n + 1$ and $(2n)^{2n} + 1$ is composite.

Proof. $n = 1$ and $n = 2$ are not solutions of the problem since $2^2 + 1$ and $4^4 + 1 = 2^8 + 1$ are primes. We will prove that all $n > 2$ are solutions. Suppose that $n > 2$ and that $n^n + 1$ and $(2n)^{2n} + 1$ are primes. By problem 1, n must be a power of 2, say $n = 2^k$. Then $n^n + 1 = 2^{k \cdot 2^k} + 1$ is a prime, hence $k \cdot 2^k$ is a power of 2 and so k is a power of 2. Next, $(2n)^{2n} + 1 = 2^{(k+1)2^{k+1}} + 1$ is prime, hence $(k+1)2^{k+1}$ is a power of 2 and so $k+1$ is a power of 2. But then k and $k+1$ are consecutive numbers and both powers of 2, thus $k = 1$ and $n = 2$, a contradiction. \square

4. For which positive integers n the numbers $2^n + 3$ and $2^n + 5$ are both primes?

Proof. It is not difficult to check that $n = 1$ and $n = 3$ are solutions, while $n = 2$ is not a solution. We claim that no $n > 3$ is a solution. Assume that $n > 3$ and that both $2^n + 3$ and $2^n + 5$ are primes. If $n \equiv 1 \pmod{3}$, then $7 \mid 2^n + 5$ and $2^n + 5 > 7$, a contradiction. If $n \equiv 2 \pmod{3}$ then $7 \mid 2^n + 3$ and $2^n + 3 > 7$, again a contradiction. Hence n is a multiple of 3. Also, n is clearly odd since otherwise $2^n + 5$ would be a multiple of 3. Thus $n \equiv 3 \pmod{6}$, say $n = 6k + 3$. If k is odd, then $2^n + 3 = 8^{2k+1} + 3$ is a multiple of 5, impossible. Hence k is even, but then $13 \mid 2^n + 5 = 8^{2k+1} + 5$ and $8^{2k+1} + 5 > 13$, again a contradiction. \square

5. (St. Petersburg 1996) Integers a, b, c have the property that the roots of the polynomial $X^3 + aX^2 + bX + c$ are pairwise relatively prime and distinct positive integers. Prove that if the polynomial $aX^2 + bX + c$ has a positive integer root, then $|a|$ is composite.

Proof. Let x_1, x_2, x_3 be the roots of the polynomial $X^3 + aX^2 + bX + c$. Then $x_1 + x_2 + x_3 = -a$, thus $|a| \geq 3$ since $x_1, x_2, x_3 \geq 1$ by assumption. If a is even, then clearly $|a|$ is composite, so assume that a is odd. Then $x_1 + x_2 + x_3$ is odd, so either x_1, x_2, x_3 are all odd, or exactly one of them is odd. This latter case is excluded by the assumption that x_1, x_2, x_3 are pairwise relatively prime. Thus x_1, x_2, x_3 are all odd. Since $b = x_1x_2 + x_2x_3 + x_3x_1$ and $-c = x_1x_2x_3$, it follows that b and c are odd. But then $ax^2 + bx + c$ cannot have integer roots, since if y is an integer root then $ay^2 + by + c \equiv y^2 + y + 1 \equiv 1 \pmod{2}$. Thus a is even, and we are done. \square

6. (Vojtech Jarník Competition 2009) Prove that if $k > 2$ then $2^{2^k-1} - 2^k - 1$ is composite.

Proof. Let $N = 2^{2^k-1} - 2^k - 1$, then

$$2N = 2^{2^k} - 1 - (2^{k+1} + 1) = (2-1)(2+1)(2^2+1)\dots(2^{2^{k-1}}+1) - (2^{k+1}+1).$$

If $k+1 = 2^m n$ with $m \geq 0$ and n odd, then $2^{k+1} + 1$ is a multiple of $2^{2^m} + 1$, and $(2+1)(2^2+1)\dots(2^{2^{k-1}}+1)$ is also a multiple of $2^{2^m} + 1$, since $m \leq k-1$ (indeed $m < 2^m \leq 2^m n = k+1$). Thus $2N$ is a multiple of $2^{2^m} + 1$ and so $2^{2^m} + 1 \mid N$. On the other hand, suppose that $N = 2^{2^m} + 1$, then since $N \equiv -1 \pmod{4}$ we must have $2^{2^m} \equiv -2 \pmod{4}$ and so $m = 0$, but then $N = 3$ which is impossible since $N > 3$. \square

7. A positive integer which is congruent 1 modulo 4 has two different representations as a sum of two squares. Prove that this number is composite.

Proof. Let n be our positive integer and consider two representations $n = x^2 + y^2 = u^2 + v^2$ as a sum of two squares. Since $n \equiv 1 \pmod{4}$, exactly one of x, y is odd, and similarly exactly one of u, v is odd. We may assume that x, u are odd and, without loss of generality, that $x > u$. Note that $\gcd(x-u, v-y)$ is then an even integer, say $2d$ for some positive integer d . Write $x - u = 2ad$ and $v - y = 2bd$ with $\gcd(a, b) = 1$. Since $(x-u)(x+u) = (v-y)(v+y)$, we easily obtain $au + a^2d = by + b^2d$. Note that this common value is divisible by a and b , thus (since $\gcd(a, b) = 1$) it is divisible by ab . Write $au + a^2d = by + b^2d = abc$ for some c . Therefore $u = bc - ad$ and $y = ac - bd$. But then $x = u + 2ad = bc + ad$ and $v = y + 2bd = ac + bd$. We finally obtain

$$n = x^2 + y^2 = (ac - bd)^2 + (bc + ad)^2 = (a^2 + b^2)(c^2 + d^2),$$

which clearly shows that n is composite. \square

Remark 8.8. By Euler's theorem (which will be discussed later on) each prime of the form $4k+1$ can be represented as a sum of two squares. Hence the problem above implies that a number $n = 4k+1$ is a prime iff it has only one representation as a sum of two squares.

8. (Moscow Olympiad) Is there an 1997-digit composite number such that if any three of its consecutive digits are replaced by any other triplet of digits then the resulting number is composite?

Proof. Such a number does exist. Let A be the product of all odd numbers from 1001 to 1997. Since each of these numbers is less than 2000 we see that

$$A < 2000^{500} = 2^{500} 10^{1500} = 32^{100} 10^{1500} < 100^{100} 10^{1500} = 10^{1700}.$$

Now we write several 0's and an 1 to the end of A and then three more 0's so that the total number of digits be equal to 1997. This number, call it N , is composite since it is even and has the desired property. Indeed, if the last digit of N is not replaced then the new number is even. If the last three 0's of N are replaced by an odd number \overline{abc} then the last four digits of the new number form the number $\overline{1abc}$ which divides N . \square

9. (AMM 10947) Prove that $\frac{5^{5n}-1}{5^n-1}$ is composite for all $n \geq 1$.

Proof. Suppose that n is even, say $n = 2k$. Then setting $x = 5^k$ we have

$$\frac{5^{5n}-1}{5^n-1} = \frac{x^{10}-1}{x^2-1} = \frac{x^5-1}{x-1} \cdot \frac{x^5+1}{x+1}.$$

Since both factors are clearly integers greater than 1, we are done.

Assume now that n is odd. The key ingredient is the identity

$$X^4 + X^3 + X^2 + X + 1 = (X^2 + 3X + 1)^2 - 5X(X + 1)^2.$$

Taking $X = 5^n$ with $n = 2k + 1$ we obtain

$$\begin{aligned} \frac{5^{5n}-1}{5^n-1} &= (5^{2n} + 3 \cdot 5^n + 1)^2 - (5^{k+1}(5^n + 1))^2 \\ &= (5^{2n} + 3 \cdot 5^n + 1 - 5^{k+1}(5^n + 1))(5^{2n} + 3 \cdot 5^n + 1 + 5^{k+1}(5^n + 1)). \end{aligned}$$

It suffices to check that

$$5^{2n} + 3 \cdot 5^n > 5^{k+1}(5^n + 1),$$

which is equivalent to $5^k(5^n + 3) > 5^n + 1$. Since this last inequality is clear, we are done. \square

10. Let $n > 1$ be an integer. Prove that the equation

$$(x+1)(x+2)\dots(x+n) = y^n$$

has no solution in positive integers.

Proof. Assume that (x, y) is a solution. Since $(x+1)(x+2)\dots(x+n)$ is between $(x+1)^n$ and $(x+n)^n$, we can write $y = x+k$ for some $k \in \{2, 3, \dots, n-1\}$. If $p \mid x+k+1$ is a prime, then by assumption $p \mid y^n$ and so $p \mid x+k$, a contradiction. The result follows. \square

11. Let n be a positive integer. Prove that if n divides $\binom{n}{k}$ for all $1 \leq k \leq n-1$, then n is prime.

Proof. Suppose on the contrary that n has a prime factor $p < n$.

By hypothesis $\frac{\binom{n}{p}}{n}$ is an integer, in other words

$$\frac{(n-1)(n-2)\dots(n-p+1)}{p!}$$

is an integer, obviously impossible (as the numerator is not a multiple of p). Hence n must be prime. \square

12. (USAMTS 2009) Find a positive integer n such that all prime factors of

$$\frac{(n+1)(n+2)\dots(n+500)}{500!}$$

are greater than 500.

Proof. The simplest way to ensure this is to choose n such that

$$\frac{(n+1)(n+2)\dots(n+500)}{500!} \equiv 1 \pmod{500!},$$

since any prime not exceeding 500 divides $500!$ and so does not divide any number congruent to 1 modulo $500!$. The previous congruence is equivalent to

$$(n+1)\dots(n+500) \equiv 500! \pmod{(500!)^2}.$$

But it is very simple to find such n 's: simply choose any multiple of $(500!)^2$. Indeed, for such n we have $n+i \equiv i \pmod{(500!)^2}$, thus

$$(n+1)\dots(n+500) \equiv 500! \pmod{(500!)^2}. \quad \square$$

13. (Russia 1999) Prove that any positive integer is the difference of two positive integers with the same number of prime factors (without counting multiplicities).

Proof. If n is even, simply write $n = 2n - n$, so assume that n is odd. If p is the smallest odd prime not dividing n (note that odd primes not dividing n certainly exist, for instance prime factors of $n+2$), then $n = pn - (p-1)n$. Since all odd prime factors of $p-1$ divide n (by minimality of p) and $p-1$ is even, pn and $(p-1)n$ have the same number of prime factors (and this is equal to the number of prime factors of n plus 1). \square

14. (Saint Petersburg) An infinite sequence $(a_n)_{n \geq 1}$ of composite numbers satisfies

$$a_{n+1} = a_n - p_n + \frac{a_n}{p_n}$$

for all n , where p_n is the smallest prime factor of a_n . If all terms of the sequence are multiples of 37, what are the possible values of a_1 ?

Proof. Since a_n and a_{n+1} are multiples of 37, so is $\frac{a_n}{p_n} - p_n$. If $p_n \neq 37$, then $\frac{a_n}{p_n}$ is a multiple of 37, while p_n is not, thus $\frac{a_n}{p_n} - p_n$ is not a multiple of 37, a contradiction. Thus $p_n = 37$ for all n . We deduce that

$$a_{n+1} - a_n = \frac{a_n}{37} - 37$$

for all n . Letting $b_n = a_n - 37^2$, we have

$$b_{n+1} = \frac{38}{37}b_n,$$

thus $b_n = \frac{38^{n-1}}{37^{n-1}}b_1$ for all $n \geq 1$. Since b_n is an integer, we deduce that $37^{n-1} \mid b_1$ for all n , which forces $b_1 = 0$ and $a_1 = 37^2$. Conversely, if $a_1 = 37^2$, then setting $a_n = 37^2$ for all n yields a sequence satisfying all conditions of the problem. \square

15. Prove that there are infinitely many pairs (a, b) of distinct positive integers a, b such that a and b have the same prime divisors, and $a + 1$ and $b + 1$ also have the same prime divisors.

Proof. Let $n \geq 2$ and let $a = 2^n - 2$ and $b = 2^n(2^n - 2)$. Then a and b clearly have the same prime divisors, and $b + 1 = (a + 1)^2$, so $a + 1$ and $b + 1$ also have the same prime divisors. \square

16. Let a, b, c, d, e, f be positive integers such that $abc = def$. Prove that $a(b^2 + c^2) + d(e^2 + f^2)$ is composite.

Proof. Suppose that $p = a(b^2 + c^2) + d(e^2 + f^2)$ is a prime. Multiplying the congruence $a(b^2 + c^2) \equiv -d(e^2 + f^2) \pmod{p}$ by ef and using the hypothesis yields

$$aef(b^2 + c^2) \equiv -abc(e^2 + f^2) \pmod{p}.$$

Note that $p > a$, so p does not divide a and so the previous congruence yields

$$ef(b^2 + c^2) + bc(e^2 + f^2) \equiv 0 \pmod{p}.$$

The left-hand side factors as $(ce + bf)(be + cf)$ and so p divides one of the numbers $ce + bf$ or $be + cf$. On the other hand

$$p = a(b^2 + c^2) + d(e^2 + f^2) \geq b^2 + c^2 + e^2 + f^2 \geq 2ce + 2bf > ce + bf$$

and similarly $p > be + cf$, a contradiction. Hence p is composite. \square

17. (Kvant M 1762) Is there a positive integer n with 2013 prime divisors such that n divides $2^n + 1$?

Proof. The answer is positive. We will prove by induction that for each $k \geq 1$ we can find n_k with exactly k prime divisors, such that $3 \mid n_k$ and $n_k \mid 2^{n_k} + 1$. If $k = 1$ take $n_1 = 3$. Assume now that $n = n_k$ is a multiple of 3, has k prime factors and satisfies $n \mid 2^n + 1$. Clearly n is odd, hence $3 \mid 2^{2n} - 2^n + 1$ and so

$$2^{3n} + 1 = (2^n + 1)(2^{2n} - 2^n + 1)$$

is a multiple of $3n$. Note that

$$2^{2n} - 2^n + 1 = (2^n - 2)(2^n + 1) + 3$$

is not divisible by 9 since $2^n - 2$ and $2^n + 1$ are both divisible by 3 for odd n . Hence the number $2^{2n} - 2^n + 1$ has a prime divisor $p > 3$. The number p is not a divisor of n since otherwise it would divide $\gcd(2^n + 1, 2^{2n} - 2^n + 1) = 3$. Hence the number $n_{k+1} = 3pn$ has $k + 1$ divisors and divides $2^{n_{k+1}} + 1$. \square

18. (Poland 2000) Let p_1 and p_2 be prime numbers and for $n \geq 3$ let p_n be the greatest prime factor of $p_{n-1} + p_{n-2} + 2000$. Prove that the sequence $(p_n)_{n \geq 1}$ is bounded.

Proof. First, observe that

$$p_n \leq \max(p_{n-1}, p_{n-2}) + 2002 \quad (*)$$

Indeed, if p_{n-1}, p_{n-2} are both odd then $p_{n-1} + p_{n-2} + 2000$ is even and greater than 2, thus

$$p_n \leq \frac{p_{n-1} + p_{n-2} + 2000}{2} < \max(p_{n-1}, p_{n-2}) + 2002,$$

while if at least one of p_{n-1}, p_{n-2} is 2 we have

$$p_n \leq p_{n-1} + p_{n-2} + 2000 \leq \max(p_{n-1}, p_{n-2}) + 2002.$$

This being established, let $M = \max(p_1, p_2) \cdot 2003! + 2$ and let us prove by induction that $p_n < M$ for all n . This is clear for $n = 1, 2$ and if it holds up to $n - 1$, then relation (*) shows that $p_n < M + 2002$. But since $M, M + 1, \dots, M + 2001$ are all composite numbers, we deduce that $p_n < M$ and we are done. \square

19. (Italy 2011) Find all primes p for which $p^2 - p - 1$ is the cube of an integer.

Proof. Clearly $p = 2$ is a solution of the problem, so assume from now on that $p > 2$. Let n be an integer such that $p^2 - p - 1 = n^3$. Then

$$p(p - 1) = n^3 + 1 = (n + 1)(n^2 - n + 1),$$

hence p divides $n + 1$ or $n^2 - n + 1$. Assume that $p \mid n + 1$, thus $n \geq p - 1$ and then $p^2 - p - 1 \geq (p - 1)^3$, which implies that $(p - 1)^3 < p(p - 1)$ and then $p^2 - 3p + 1 < 0$, impossible since $p \geq 3$.

Hence $p \mid n^2 - n + 1$, say $n^2 - n + 1 = kp$ for some positive integer k . Coming back to the relation $p(p - 1) = (n + 1)(n^2 - n + 1)$ yields $p - 1 = k(n + 1)$, hence

$$n^2 - n + 1 = kp = k(1 + k(n + 1)) = k + k^2(n + 1).$$

This can be rewritten as

$$n^2 - (1 + k^2)n + 1 - k - k^2 = 0.$$

Considering this as a quadratic equation in n , its discriminant

$$\Delta = (1 + k^2)^2 + 4(k^2 + k - 1)$$

must be a square, since the equation has an integer root. One easily checks that Δ is not a square for $k \leq 2$, and it is a square for $k = 3$, in which case $n^2 - 10n - 11 = 0$ yields $n = 11$ and then $p = 37$. Assume now that $k > 3$. Then an easy computation shows that

$$\Delta = (k^2 + 3)^2 + 4(k - 3) > (k^2 + 3)^2$$

and since Δ is a square we must have $\Delta \geq (k^2 + 4)^2$, which yields

$$4(k - 3) \geq (k^2 + 4)^2 - (k^2 + 3)^2 = 2k^2 + 7.$$

This last inequality is impossible for $k > 3$, hence the only solutions of the problem are $p = 2$ and $p = 37$. \square

Remark 8.9. A similar problem (with $p^2 - p + 1$ instead of $p^2 - p - 1$) was proposed in Saint Petersburg in 1995 and later on at the Balkan Mathematical Olympiad in 2005. The solution to this new problem is $p = 19$. Yet another similar problem was proposed at the Tuymaada Olympiad in 2013: find all primes p, q such that $p^2 - pq - q^3 = 1$.

20. (Kvant M 2145) Let $x > 2, y > 1$ be integers such that $x^y + 1$ is a perfect square. Prove that x has at least 3 different prime divisors.

Proof. Write $x^y + 1 = a^2$ for a positive integer a . Assume first that x is a power of a prime. Thus $(a - 1)(a + 1)$ is a power of a prime, in particular both $a - 1$ and $a + 1$ are powers of that prime, and both are greater than 1. Since they differ by 2, the prime must be 2 and $a - 1 = 2$, thus $a = 3$ and $x^y = 8$, contradicting the fact that $x > 2$ and $y > 1$.

Assume now that x has precisely two prime factors, say $p < q$. We have $(a - 1)(a + 1) = x^y$. If $\gcd(a - 1, a + 1) = 1$, then $a - 1$ and $a + 1$ must be y th powers, say $a - 1 = b^y$ and $a + 1 = c^y$, so that $c^y - b^y = 2$. This is impossible, since

$$c^y - b^y = (c - b)(c^{y-1} + \dots + b^{y-1}) \geq 2^{y-1} + 1 \geq 3.$$

Thus $\gcd(a-1, a+1)$ is not 1, and since it divides 2 it must be equal to 2. In particular $p = 2$. Since $(a-1)(a+1) = x^y$ and the prime factors of x are 2 and q , we have two possibilities:

- i) Either $a-1 = 2q^{uy}$ and $a+1 = 2^{vy-1}$ for some integers u, v . But then $2^{vy-2} - 1 = q^{uy}$, contradicting lemma 8.10 below.
- ii) Or $a-1 = 2^{uy-1}$ and $a+1 = 2q^{vy}$ for some integers u, v . Then $2^{uy-2} + 1 = q^{vy}$. Using again lemma 8.10 below, we obtain $uy - 2 = 3$ and $vy = 2$, impossible. This finishes the proof. \square

Lemma 8.10. a) $2^n - 1$ is not a perfect power if $n > 1$.

b) $2^n + 1$ is a perfect power only for $n = 3$.

Proof. a) Suppose that $2^n - 1 = a^b$, with $a, b > 1$. Since $2^n - 1$ is of the form $4k + 3$, it cannot be a square, so b is odd. Then

$$2^n = (1+a)(1-a+a^2-\dots+a^{b-1}).$$

Thus $1+a$ and $1-a+\dots+a^{b-1}$ are powers of two. The second number is odd, since a and b are. Thus we must have $1-a+\dots+a^{b-1} = 1$ and $2^n = 1+a$. This yields $1+a = 1+a^b$, contradicting the inequality $a^b > a$.

b) Clearly $2^1 + 1 = 3$ and $2^2 + 1 = 5$ are not perfect powers. Assume that $2^n + 1$ is a perfect power for some $n > 3$, say $2^n + 1 = x^k$ for some $x, k > 1$. Then x is odd. If k is odd, then $1+x+\dots+x^{k-1}$ is odd, greater than 1 and it divides 2^n , a contradiction. Hence k is even, say $k = 2l$. Then $(x^l - 1)(x^l + 1) = 2^n$ and so $x^l - 1$ and $x^l + 1$ are powers of 2 differing by 2. This forces $x^l - 1 = 2$, then $x = 3$ and $l = 1$, that is $k = 2$, and finally $n = 3$. \square

Remark 8.11. Using more advanced techniques (the Birkhoff-Vandiver theorem) one can prove that x has at least $1 + \tau(y)$ prime divisors.

21. (Russia 2010) Prove that for any $n > 1$ there are n consecutive positive integers whose product is divisible by all primes not exceeding $2n + 1$, and not divisible by any other prime.

Proof. All prime factors of $(n+2)(n+3)\dots(2n+1)$ are less than or equal to $2n+1$. On the other hand, $(n+2)\dots(2n+1)$ is a multiple of $n!$, since it is the product of n consecutive integers. Thus if $n+1$ is not a prime, then $(n+2)\dots(2n+1)$ is divisible by all primes not exceeding $2n+1$. Assume that $n+1$ is a prime, then $n+2$ is not a prime and $(n+3)\dots(2n+1)(2n+2)$ is divisible by $n!$ (for the same reason as above) and by $n+1, n+3, \dots, 2n+1$. Since $n+2$ is not a prime, we deduce that $(n+3)\dots(2n+1)(2n+2)$ is divisible exactly by the prime numbers not exceeding $2n+1$, and the problem is solved in all cases. \square

22. (Iran 2015) Prove that infinitely many positive integers n cannot be written as the sum of two positive integers all of whose prime factors are less than 1394.

Proof. Let p_1, \dots, p_k be all primes not exceeding 1394 and let S_n be the set of numbers $j \in \{1, 2, 3, \dots, 2^n\}$ all of whose prime factors are among p_1, \dots, p_k . Any such number j is of the form $p_1^{\alpha_1} \dots p_k^{\alpha_k}$ for a unique k -tuple of nonnegative integers $\alpha_1, \dots, \alpha_k$. Since $p_i \geq 2$ and $j \leq 2^n$, we must have $2^{\alpha_i} \leq 2^n$ for all i , thus $\alpha_i \leq n$ for all i . It follows that there are at most $(1+n)^k$ such k -tuples and so

$$|S_n| \leq (1+n)^k.$$

It follows that there are at most $(n+1)^{2k}$ numbers between 1 and 2^n that can be written as the sum of two numbers in S_n . If n is large enough, then $(1+n)^{2k} < \frac{1}{2}2^n$ (note that by the binomial formula $2^n > \binom{n}{2k+1}$ if $n > 2k+1$, and $\binom{n}{2k+1}$ is a polynomial expression of degree $2k+1$ in n). Thus for n large enough more than half of the numbers between 1 and 2^n are solutions of the problem, yielding the result. Note that the proof can be interpreted as saying that the probability that an integer is a sum of two numbers whose prime factors are $\leq k$ is 0. \square

23. (China 2007) Let $n > 1$ be an integer. Prove that $2n-1$ is a prime number if and only if for any n pairwise distinct positive integers

a_1, a_2, \dots, a_n there exist $i, j \in \{1, 2, \dots, n\}$ such that

$$\frac{a_i + a_j}{\gcd(a_i, a_j)} \geq 2n - 1$$

Proof. Suppose first that $p = 2n - 1$ is a prime and let a_1, \dots, a_n be pairwise distinct positive integers. Suppose that

$$\frac{a_i + a_j}{\gcd(a_i, a_j)} < p$$

for $i, j \in \{1, \dots, n\}$. Dividing each of the numbers a_1, \dots, a_n by $\gcd(a_1, \dots, a_n)$, we may assume that $\gcd(a_1, \dots, a_n) = 1$.

If there is i such that $p \mid a_i$, then we can choose j such that p does not divide a_j and then p does not divide $\gcd(a_i, a_j)$. Thus p divides $\frac{a_i}{\gcd(a_i, a_j)}$ and we obtain the plain contradiction

$$p \leq \frac{a_i}{\gcd(a_i, a_j)} < \frac{a_i + a_j}{\gcd(a_i, a_j)} < p.$$

Suppose now that a_1, a_2, \dots, a_n are not multiples of p . By the pigeonhole principle, two of the numbers $a_1, a_2, \dots, a_n, -a_1, \dots, -a_n$ must give the same remainder when divided by p . So we can find $i \neq j$ such that $p \mid a_i + a_j$ or $p \mid a_i - a_j$. Note that p does not divide $\gcd(a_i, a_j)$, so p divides $\frac{a_i \pm a_j}{\gcd(a_i, a_j)}$ for a suitable choice of the sign \pm . We obtain again a contradiction

$$p > \frac{a_i + a_j}{\gcd(a_i, a_j)} \geq \left| \frac{a_i \pm a_j}{\gcd(a_i, a_j)} \right| \geq p.$$

So our initial assumption was wrong and the result follows.

Suppose now that $2n - 1$ is composite, so we can write it as xy , with $x, y > 1$. Define n integers a_1, a_2, \dots, a_n by choosing the first x positive integers $1, 2, \dots, x$, then the next $n - x$ even numbers $x + 1, x + 3, \dots, xy - x$. It is not difficult to check that $\frac{a_i + a_j}{\gcd(a_i, a_j)} < 2n - 1$ for all i, j . \square

24. (Tournament of the Towns 2009) Initially the number 6 is written on a blackboard. At the n th step, one replaces the number d written on the blackboard with $d + \gcd(d, n)$. Prove that at each step the number on the blackboard increases either by 1 or by a prime number.

Proof. This problem is very difficult! Let a_n be the number on the blackboard at step n , so that $a_0 = 6$ and

$$a_n = a_{n-1} + \gcd(a_{n-1}, n).$$

Let $b_n = a_n - a_{n-1}$, hence we need to prove that b_n is either 1 or a prime for all n . The first few values of the sequence b_1, b_2, \dots are 1, 1, 1, 1, 5, 3, 1, 1, 1, 1, 11, 3,

The crucial claim is the following: suppose that $a_n = 3n$ and that $b_{n+1} = 1$. Let k be the smallest positive integer such that $b_{n+k} \neq 1$. Then b_{n+k} is a prime and $a_{n+k} = 3(n+k)$. We will prove this claim by induction. It is not difficult to check it for $n \leq 5$ using the previous explicit values for the sequence $(b_n)_{n \geq 1}$. It is not difficult to see that $a_{n+1} = 3n + 1$, $a_{n+2} = 3n + 2, \dots$, $a_{n+k-1} = 3n + k - 1$ and so

$$b_{n+k} = \gcd(n+k, 3n+k-1) = \gcd(n+k, 2k+1) \mid 2k+1.$$

Suppose that $2k+1$ is not a prime and let p be a prime factor of $\gcd(2k+1, n+k)$. Then $p \leq \frac{2k+1}{3} < k$ and

$$b_{n+k-p} = \gcd(n+k-p, 3n+k-p-1) = \gcd(n+k-p, 2k+1-2p)$$

is a multiple of p , contradicting the fact that k was minimal with $b_{n+k} \neq 1$. Thus $2k+1$ is a prime and $b_{n+k} = 2k+1$, hence

$$a_{n+k} = a_{n+k-1} + b_{n+k} = 3n + k - 1 + 2k + 1 = 3(n+k),$$

finishing the induction. It is clear that the claim implies the desired result. \square

25. (Komal) Is it possible to find 2000 positive integers such that none of them is divisible by any of the other numbers but the square of each is divisible by all the others?

Proof. The answer is positive. Let $k = 2000$, p_1, \dots, p_k pairwise distinct primes and let $P = p_1 \dots p_k$ and $x_i = \frac{P^2}{p_i}$ for $1 \leq i \leq k$. Then x_1, \dots, x_k are positive integers, $\frac{x_i}{x_j} = \frac{p_j}{p_i}$ is not an integer if $i \neq j$, yet $x_i^2 = \frac{P^2}{p_i^2} \cdot P^2$ is a multiple of P^2 , which is a multiple of any x_j , with $1 \leq j \leq k$. The result follows. \square

26. A positive integer n is called powerful if $p^2 \mid n$ for any prime factor p of n . Prove that there are infinitely many pairs of consecutive powerful numbers.

Proof. The key observation is that if n and $n+1$ are powerful, then so are $4n(n+1)$ and $4n(n+1)+1 = (2n+1)^2$. This is clear by the definition of powerful numbers. Since 8 and 9 are powerful, the result follows. \square

27. Let p_n be the largest prime not exceeding n and let q_n be the smallest prime larger than n . Prove that for all $n > 1$ we have

$$\sum_{k=2}^n \frac{1}{p_k q_k} < \frac{1}{2}.$$

Proof. Let r_1, r_2, \dots be the increasing sequence of primes and write $q_n = r_m$ for some positive integer m . Since $p_k = r_i$ and $q_k = r_{i+1}$ for $p_i \leq k < p_{i+1}$, it follows that

$$\begin{aligned} \sum_{k=2}^n \frac{1}{p_k q_k} &\leq \sum_{k=2}^{p_m-1} \frac{1}{p_k q_k} = \sum_{i=1}^{m-1} \sum_{k=r_i}^{r_{i+1}-1} \frac{1}{r_i r_{i+1}} = \sum_{i=1}^{m-1} \frac{r_{i+1} - r_i}{r_i r_{i+1}} \\ &= \sum_{i=1}^{m-1} \left(\frac{1}{r_i} - \frac{1}{r_{i+1}} \right) = \frac{1}{2} - \frac{1}{r_m} < \frac{1}{2}. \end{aligned} \quad \square$$

28. (Russia 2010) Are there infinitely many positive integers which cannot be expressed as $\frac{x^2-1}{y^2-1}$, with x, y integers greater than 1?

Proof. We will prove that p^2 cannot be expressed as $\frac{x^2-1}{y^2-1}$ for any odd prime p , thus the answer is positive. Assume that $\frac{x^2-1}{y^2-1} = p^2$, that is $x^2 - 1 = p^2(y^2 - 1)$. Then $p \mid (x-1)(x+1)$, hence $p \mid x-1$ or $p \mid x+1$. Moreover, we have $\gcd(x-1, x+1) = 2$, thus necessarily $p^2 \mid x-1$ or $p^2 \mid x+1$. Assume that $p^2 \mid x-1$, say $x-1 = kp^2$, then $k(kp^2+2) = y^2-1$, or equivalently

$$(kp)^2 + 2k + 1 = y^2.$$

On the other hand

$$(kp)^2 + 2k + 1 < (kp)^2 + 2kp + 1 = (kp+1)^2,$$

hence $kp < y < kp+1$, a contradiction. Similarly, if $p^2 \mid x+1$, say $x = kp^2 - 1$ then $(kp)^2 - 2k + 1 = y^2$ and

$$(kp)^2 - 2k + 1 > (kp)^2 - 2kp + 1 = (kp-1)^2,$$

hence $kp-1 < y < kp$, a contradiction. \square

29. (Baltic Way 2004) Is there an infinite sequence of prime numbers p_1, p_2, \dots such that $|p_{n+1} - 2p_n| = 1$ for each $n \geq 1$?

Proof. Suppose that such a sequence exists and suppose that there is i such that $p_i > 3$. Suppose that $p_i \equiv 1 \pmod{3}$, then $2p_i+1$ is a multiple of 3 and greater than 3, thus necessarily $p_{i+1} = 2p_i - 1 \equiv 1 \pmod{3}$. Repeating the argument yields $p_{i+k} = 2p_{i+k-1} - 1$ for $k \geq 1$, then by induction $p_{i+k} = 2^k p_i - 2^k + 1$. Thus $2^k p_i - 2^k + 1$ is a prime for all $k \geq 1$. Since p_i is odd, there is $k > 0$ such that $p_i \mid 2^k - 1$, then $p_i \mid 2^k p_i - 2^k + 1$ and so $p_i = 2^k p_i - 2^k + 1$, that is $(p_i - 1)(2^k - 1) = 0$. This is absurd, so we must have $p_i \equiv -1 \pmod{3}$. Then $2p_i - 1$ is a multiple of 3 greater than 3, hence $p_{i+1} = 2p_i + 1 \equiv -1 \pmod{3}$. Repeating the above arguments, we deduce that $p_{i+k} = 2^k p_i + 2^k - 1$ for $k \geq 1$. Choosing $k \geq 1$ such that $p_i \mid 2^k - 1$ (which is possible by corollary 4.15) yields a contradiction. We deduce that $p_i \leq 3$ for all i , and this is obviously impossible. Thus there is no such sequence. \square

30. Let a_1, a_2, \dots, a_k be positive real numbers such that for all but finitely many positive integers n we have

$$\gcd(n, \lfloor a_1 n \rfloor + \lfloor a_2 n \rfloor + \dots + \lfloor a_k n \rfloor) > 1.$$

Prove that a_1, \dots, a_k are integers.

Proof. Let N be a positive integer such that for $n > N$ we have

$$\gcd(n, \lfloor a_1 n \rfloor + \lfloor a_2 n \rfloor + \dots + \lfloor a_k n \rfloor) > 1.$$

Let p_1, p_2, \dots be the sequence of primes greater than N , then for all $i \geq 1$ the quotient

$$x_i = \frac{\lfloor a_1 p_i \rfloor + \lfloor a_2 p_i \rfloor + \dots + \lfloor a_k p_i \rfloor}{p_i}$$

is an integer. On the other hand, since $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$ for all x , we have

$$a_1 + \dots + a_k - \frac{k}{p_i} < x_i \leq a_1 + \dots + a_k.$$

Since this happens for all i , it is not difficult to deduce that $a_1 + \dots + a_k$ is an integer and $x_i = a_1 + \dots + a_k$ for all sufficiently large i , say $i > i_0$. But then

$$\{a_1 p_i\} + \dots + \{a_k p_i\} = 0$$

for $i > i_0$, where $\{x\}$ is the fractional part of x . This forces $a_j p_i \in \mathbf{Z}$ for $1 \leq j \leq k$ and $i > i_0$. Using Bézout's theorem, this immediately implies that a_1, \dots, a_k are all integers. \square

31. (IMO Shortlist 2006) We define a sequence a_1, a_2, a_3, \dots by setting

$$a_n = \frac{1}{n} \left(\left\lfloor \frac{n}{1} \right\rfloor + \left\lfloor \frac{n}{2} \right\rfloor + \dots + \left\lfloor \frac{n}{n} \right\rfloor \right)$$

for every positive integer n .

- a) Prove that $a_{n+1} > a_n$ for infinitely many n .
b) Prove that $a_{n+1} < a_n$ for infinitely many n .

Proof. a) Assuming the opposite, it follows that the sequence $(a_n)_n$ is bounded (since if $a_{n+1} \leq a_n$ for $n \geq N$, then $a_n \leq \max(a_1, \dots, a_N)$ for all n). However

$$a_n > \frac{1}{n} \left(\frac{n}{1} - 1 + \frac{n}{2} - 1 + \dots + \frac{n}{n} - 1 \right) = 1 + \frac{1}{2} + \dots + \frac{1}{n} - 1$$

and the last expression is not bounded. This contradiction settles part a).

b) Note that $a_{n+1} < a_n$ is equivalent to

$$\sum_{k=1}^{n+1} \left\lfloor \frac{n+1}{k} \right\rfloor < \left(1 + \frac{1}{n} \right) \sum_{k=1}^n \left\lfloor \frac{n}{k} \right\rfloor$$

or equivalently

$$1 + \sum_{k=1}^n \left(\left\lfloor \frac{n+1}{k} \right\rfloor - \left\lfloor \frac{n}{k} \right\rfloor \right) < \frac{1}{n} \sum_{k=1}^n \left\lfloor \frac{n}{k} \right\rfloor = a_n.$$

The key observation is that $\left\lfloor \frac{n+1}{k} \right\rfloor - \left\lfloor \frac{n}{k} \right\rfloor$ equals 0 if k does not divide $n+1$ and 1 otherwise. This is a simple exercise using the Euclidean division that we leave to the reader. Therefore we can rewrite the previous inequality as

$$1 + \sum_{k \leq n, k|n+1} 1 < a_n.$$

This suggests taking $n = p - 1$ with p a prime, so that the left-hand side is extremely simple: it reduces to 2. So it suffices to prove that $a_{p-1} > 2$ for infinitely many primes p , which is the case, since we have already seen in part a) that a_n tends to ∞ . \square

32. (APMO 1994) Find all integers n of the form $a^2 + b^2$ with a, b relatively prime positive integers, such that any prime $p \leq \sqrt{n}$ divides ab .

Proof. If $p \leq \sqrt{n}$ then p divides a or b . Since $\gcd(a, b) = 1$ we have $\gcd(a, a^2 + b^2) = \gcd(b, a^2 + b^2) = 1$ and so p does not divide n , which

implies that n is a prime number. Next, let p_1, \dots, p_k be all primes less than \sqrt{n} . Then $p_{k+1} > \sqrt{n}$. Assume that $k \geq 4$, then Bonse's inequality yields

$$ab \geq p_1 \dots p_k > p_{k+1}^2 > n = a^2 + b^2,$$

a contradiction. Thus $k \leq 3$ and so $\sqrt{n} < 7$, that is $n < 49$. If $n \geq 25$, then $k = 3$ and $30 = p_1 p_2 p_3$ divides ab , thus $n = a^2 + b^2 \geq 2ab \geq 60$, a contradiction. Hence $n \leq 24$ and n is a prime. If $n > 9$ then $k = 2$ and $6 \mid ab$, which easily implies that one of a, b is 3 (otherwise $n > 24$) and then a direct check yields $n = 13$. If $n \leq 8$ then we want n to be a prime and $2 \mid ab$, which gives $n = 2$ or 5. \square

33. (Iran TST 2009) Find all polynomials f with integer coefficients having the following property: for all primes p and for all integers a, b , if $p \mid ab - 1$, then $p \mid f(a)f(b) - 1$.

Proof. Let a be a positive integer and let $p > a$ be a prime. Then a and p are relatively prime, so there is an integer b such that $p \mid ab - 1$. By hypothesis $f(a)f(b) \equiv 1 \pmod{p}$. Let $f(X) = a_0 + a_1X + \dots + a_nX^n$ for some integers a_0, \dots, a_n with $a_n \neq 0$. Then $ab \equiv 1 \pmod{p}$ and $a^n f(a)f(b) \equiv a^n \pmod{p}$. But

$$\begin{aligned} a^n f(b) &\equiv a_n(ab)^n + a_{n-1}(ab)^{n-1}a + \dots + a_0a^n \\ &\equiv a_n + a_{n-1}a + \dots + a_0a^n \pmod{p}. \end{aligned}$$

Hence letting $g(X) = a_n + a_{n-1}X + \dots + a_0X^n$ we obtain

$$f(a)g(a) \equiv a^n \pmod{p}.$$

Thus infinitely many primes divide $f(a)g(a) - a^n$ and so $f(a)g(a) = a^n$ for any positive integer a . It follows that $f(X)g(X) = X^n$ and so $f(X) = \pm X^d$ for some $0 \leq d \leq n$. Conversely, any polynomial $f(X) = \pm X^d$ with $d \geq 0$ is a solution of the problem. \square

34. Prove that there is a positive integer n such that the interval $[n^2, (n+1)^2]$ contains at least 2016 primes.

Proof. Let $k = 2015$ and assume that for all n there are at most k primes between n^2 and $(n+1)^2$. Pick any $N > 1$ and observe that

$$\sum_{p < N^2} \frac{1}{p} = \sum_{p < 2^2} \frac{1}{p} + \sum_{2^2 \leq p < 3^2} \frac{1}{p} + \dots + \sum_{(N-1)^2 \leq p < N^2} \frac{1}{p}.$$

By assumption each of the sums

$$\sum_{j^2 \leq p < (j+1)^2} \frac{1}{p}$$

has at most k terms, each smaller than or equal to $\frac{1}{j^2}$, thus the whole sum is bounded from above by $\frac{k}{j^2}$. We deduce that

$$\sum_{p < N^2} \frac{1}{p} \leq \sum_{j=1}^N \frac{k}{j^2} < k + k \sum_{j=2}^N \frac{1}{j(j-1)} < 2k.$$

We know however (see theorem 4.74) that for N large enough we have

$$\sum_{p < N^2} \frac{1}{p} > 2k.$$

This contradiction shows that our original assumption was wrong and the result follows. \square

35. (IMO 1977) Let $n > 2$ be an integer and let V_n be the set of integers of the form $1 + kn$ with $k \geq 1$. A number $m \in V_n$ is called indecomposable if it cannot be written as the product of two elements of V_n . Prove that there is $r \in V_n$ that can be expressed as the product of indecomposable elements of V_n in more than one way (expressions which differ only in order of the elements of V_n will be considered the same).

Proof. We have already seen (see example 4.56) that there are infinitely many primes p not congruent to 1 modulo n . Their remainders modulo n lie in a finite set, thus we can find two such primes $p, q > n$ which are congruent modulo n . Let d be the smallest positive integer such that

$p^d \equiv 1 \pmod{n}$ (it exists, thanks to corollary 4.15). Then p^d, q^d, pq^{d-1} and $p^{d-1}q$ are all indecomposable elements of V_n . Indeed, it is clear that they are in V_n (i.e. that they are congruent to 1 modulo n), and that their proper divisors are not in V_n (by minimality of d and the fact that $p \equiv q \pmod{n}$). In order to finish the proof, it suffices to observe that

$$p^d \cdot q^d = (pq^{d-1}) \cdot (qp^{d-1}). \quad \square$$

36. (German TST 2009) The sequence $(a_n)_{n \in \mathbb{N}}$ is defined by $a_1 = 1$ and

$$a_{n+1} = a_n^4 - a_n^3 + 2a_n^2 + 1$$

for all $n \geq 1$. Prove that there are infinitely many primes which do not divide any of the numbers a_1, a_2, \dots

Proof. The key ingredient in this problem is the study of the sequence $b_n = a_n^2 + 1$. Note that

$$a_{n+1} = (a_n^2 + 1)^2 - a_n^3 = b_n^2 - a_n(b_n - 1).$$

It follows that $a_{n+1} \equiv a_n \pmod{b_n}$ and so $a_{n+1}^2 + 1 \equiv a_n^2 + 1 \equiv 0 \pmod{b_n}$. In other words, b_n divides b_{n+1} for all n . We can actually refine this observation: we have

$$a_{n+1}^2 + 1 \equiv a_n^2(b_n - 1)^2 + 1 \equiv a_n^2(1 - 2b_n) + 1 \equiv b_n(1 - 2a_n^2) \pmod{b_n^2}.$$

Note that $\gcd(1 - 2a_n^2, b_n) = 1$, since any prime dividing $1 - 2a_n^2$ and $b_n = a_n^2 + 1$ would also divide $1 - 2a_n^2 + 2(a_n^2 + 1) = 3$, but 3 does not divide $a_n^2 + 1$. We conclude that $b_{n+1} = b_n c_n$ with $\gcd(b_n, c_n) = 1$. Note that clearly $a_{n+1} > a_n$ for all n , thus $c_n > 1$ for all n . Let p_k be an arbitrary prime factor of c_k , then p_k does not divide b_k (as $\gcd(b_k, c_k) = 1$) and so it does not divide $b_1 b_2 \dots b_k$ (since $b_1 \mid b_2 \mid \dots \mid b_k$). In particular p_k does not divide $c_1 c_2 \dots c_{k-1}$ and so the sequence p_1, p_2, \dots consists of pairwise distinct primes. We will prove that any of these primes is a solution of the problem.

Assume that $p \mid b_n$ for some $n \geq 1$, and that $p \mid a_k$ for some $k \geq 1$. Note that for all $n \geq 1$ we have $a_{n+1} \equiv 1 \equiv a_1 \pmod{a_n}$ and then $a_{n+2} \equiv a_1^4 - a_1^3 + 2a_1^2 + 1 \equiv a_2 \pmod{a_n}$. An immediate inductive argument shows that $a_{n+j} \equiv a_j \pmod{a_n}$ for all $n, j \geq 1$. In particular, $a_k \mid a_{jk}$ for all $j \geq 1$. Choose j such that $jk \geq n$, then $p \mid a_k \mid a_{jk}$ and so p does not divide $b_{jk} = a_{jk}^2 + 1$. This is however impossible, since $p \mid b_n \mid b_{jk}$. \square

37. Prove that for all $n \geq 1$ we have

$$\sum_{d|n} \sigma(d) = n \cdot \sum_{d|n} \frac{\tau(d)}{d}, \quad n \cdot \sum_{d|n} \frac{\sigma(d)}{d} = \sum_{d|n} d\tau(d).$$

Proof. Let us prove the first equality. Since both sides define multiplicative functions of n , it suffices to prove that they agree on prime powers, thus we may assume that $n = p^k$ for a prime p and some $k \geq 0$. Then

$$\sum_{d|n} \sigma(d) = \sum_{i=0}^k \sigma(p^i) = 1 + (1+p) + \dots + (1+p+\dots+p^k) = (k+1) + kp + \dots + p^k$$

and

$$n \cdot \sum_{d|n} \frac{\tau(d)}{d} = p^k \sum_{i=0}^k (i+1)p^{-i} = \sum_{i=0}^k p^{k-i}(i+1) = (k+1) + kp + \dots + p^k,$$

thus the two sides agree.

For the second equality we proceed similarly, reducing to the case $n = p^k$ and then computing

$$\begin{aligned} n \cdot \sum_{d|n} \frac{\sigma(d)}{d} &= p^k \sum_{i=0}^k \frac{\sigma(p^i)}{p^i} = p^k + (p^k + p^{k-1}) + \dots + (p^k + p^{k-1} + \dots + 1) \\ &= (k+1)p^k + kp^{k-1} + \dots + 1 = \sum_{i=0}^k (i+1)p^i = \sum_{d|n} \tau(d). \end{aligned}$$

Here is also an alternative solution, suggested by Richard Stong and using the convolution product of arithmetic functions. Let 1 denote the constant function with value 1 and id the identity function. We already saw that $1 * 1 = \tau$ and $1 * \text{id} = \sigma$. We easily compute that

$$(\text{id} * \text{id})(n) = \sum_{d|n} d \cdot \frac{n}{d} = n\tau(n).$$

Now the first equality just reads

$$1 * \sigma = 1 * (1 * \text{id}) = (1 * 1) * \text{id} = \tau * \text{id},$$

and the second reads

$$\sigma * \text{id} = (1 * \text{id}) * \text{id} = 1 * (\text{id} * \text{id}). \quad \square$$

38. a) Let f be a multiplicative function with $f(1) = 1$ (this is equivalent to f being nonzero). Prove that for all $n > 1$ we have

$$\sum_{d|n} f(d)\mu(d) = \prod_{p|n} (1 - f(p)),$$

the product being taken over the prime divisors of n .

- b) Deduce closed formulae for

$$\sum_{d|n} \mu(d)\tau(d), \quad \sum_{d|n} \mu(d)\sigma(d) \quad \text{and} \quad \sum_{d|n} \mu(d)\varphi(d) \quad \text{for } n > 1.$$

Proof. a) Let p_1, p_2, \dots, p_k be the distinct prime divisors of n . The only divisors d of n for which $f(d)\mu(d) \neq 0$ are products of distinct elements of the set $\{p_1, \dots, p_k\}$ (including the empty product, which equals 1 by convention). Hence

$$\sum_{d|n} f(d)\mu(d) = 1 - \sum_{i=1}^k f(p_i) + \sum_{1 \leq i < j \leq k} f(p_i p_j) + \dots + (-1)^{k-1} f(p_1 \dots p_k).$$

Since f is multiplicative, the right-hand side can further be written as

$$\begin{aligned} 1 - \sum_{i=1}^k f(p_i) + \sum_{1 \leq i < j \leq k} f(p_i)f(p_j) + \dots + (-1)^{k-1} f(p_1)\dots f(p_k) \\ = (1 - f(p_1))\dots(1 - f(p_k)). \end{aligned}$$

The result follows.

b) By using a), we obtain

$$\sum_{d|n} \mu(d)\tau(d) = \prod_{p|n} (1 - \tau(p)) = (-1)^{\omega(n)},$$

where $\omega(n)$ is the number of prime factors of n . Similarly, we obtain

$$\sum_{d|n} \mu(d)\sigma(d) = \prod_{p|n} (1 - (1 + p)) = (-1)^{\omega(n)} \cdot \prod_{p|n} p$$

and

$$\sum_{d|n} \mu(d)\varphi(d) = \prod_{p|n} (1 - (p - 1)) = \prod_{p|n} (2 - p). \quad \square$$

39. Let f be an arithmetic function such that the function g defined by

$$g(n) = \sum_{d|n} f(d)$$

is multiplicative. Prove that f is multiplicative.

Proof. By the Möbius inversion formula

$$f(n) = \sum_{d|n} \mu(d)g\left(\frac{n}{d}\right),$$

hence f is the convolution product of the multiplicative functions μ and g . Theorem 4.99 implies that f is multiplicative. \square

40. a) Let f be an arithmetic function and let g be the arithmetic function defined by

$$g(n) = \sum_{d|n} f(d).$$

For all $n \geq 1$ we have

$$\sum_{k=1}^n g(k) = \sum_{k=1}^n f(k) \left[\frac{n}{k} \right].$$

- b) Prove that the following relations hold for all $n \geq 1$

$$\sum_{k=1}^n \tau(k) = \sum_{k=1}^n \left[\frac{n}{k} \right], \quad \sum_{k=1}^n \sigma(k) = \sum_{k=1}^n k \left[\frac{n}{k} \right].$$

Proof. a) Taking into account that there are $\left[\frac{n}{k} \right]$ multiples of k in the set $\{1, 2, \dots, n\}$, we can write

$$\sum_{k=1}^n g(k) = \sum_{k=1}^n \sum_{d|n} f(d) = \sum_{d \leq n} f(d) \cdot \sum_{d|k, k \leq n} 1 = \sum_{k=1}^n f(k) \left[\frac{n}{k} \right].$$

- b) The first formula follows from the proposition by taking f the constant map 1 (so that $g(n) = \tau(n)$). For the second formula, take $f(n) = n$ in the proposition (so $g(n) = \sigma(n)$). \square

41. Let $f(n)$ be the difference between the number of positive divisors of n of the form $3k+1$ and the number of positive divisors of the form $3k-1$. Prove that f is multiplicative.

Proof. Let m, n be relatively prime positive integers. Then each positive divisor d of mn can be uniquely written as the product $d = ef$ of a positive divisor e of m and a positive divisor f of n . We have $d \equiv 1 \pmod{3}$ if and only if $e \equiv f \equiv 1 \pmod{3}$ or $e \equiv f \equiv 2 \pmod{3}$. Thus, if $g(n)$ (respectively $h(n)$) is the number of positive divisors of the form $3k+1$ (respectively $3k-1$) of n , then

$$g(mn) = g(m)g(n) + h(m)h(n).$$

Similarly, we obtain

$$h(mn) = g(m)h(n) + g(n)h(m).$$

We deduce that

$$\begin{aligned} f(mn) &= g(mn) - h(mn) = g(m)(g(n) - h(n)) - h(m)(g(n) - h(n)) \\ &= f(n)f(m), \end{aligned}$$

proving that f is multiplicative. \square

Remark 8.12. a) Once we know that f is multiplicative, it is not difficult to check that $f(n) \geq 0$ for all n . Indeed, if $p \equiv 1 \pmod{3}$ then clearly $f(p^n) = 1 + n$, while if $p \equiv 2 \pmod{3}$, then $f(p^n)$ equals 1 if n is even and 0 otherwise.

b) One can prove that the equation $x^2 - xy + y^2 = n$ has exactly $6f(n)$ solutions in integers.

c) Similarly, one can prove that for any $k \in \{4, 6, 8, 12, 24\}$ any positive integer n has at least as many positive divisors of the form $mk + 1$ as positive divisors of the form $mk - 1$. Moreover, this property does not hold for any other k .

42. (AMM 2001) Find all totally multiplicative functions $f : \mathbf{N} \rightarrow \mathbf{C}$ such that the function

$$F(n) = \sum_{k=1}^n f(k)$$

is also totally multiplicative.

Proof. There are three such functions: the functions that are identically 0, respectively 1, and the function f such that $f(1) = 1$ and $f(n) = 0$ for $n \geq 2$. For $k > 1$, we have $f(2k) = f(2)f(k)$ and

$$\begin{aligned} f(2k-1) &= F(2k) - F(2k-2) - f(2k) \\ &= F(2)(F(k) - F(k-1)) - f(2k) \\ &= (1 + f(2))f(k) - f(2)f(k) = f(k). \end{aligned}$$

Therefore, each value $f(n)$ is a power of $f(2)$. Furthermore,

$$f(2) = f(3) = f(5) = f(9) = f(3)^2 = f(2)^2.$$

Thus $f(2) \in \{0, 1\}$, and the result follows. \square

43. Find all nonzero totally multiplicative functions $f : \mathbf{N} \rightarrow \mathbf{R}$ such that $f(n+1) \geq f(n)$ for all n .

Proof. Clearly for any nonnegative real number k the function $f(n) = n^k$ is a solution of the problem. We will prove that these are all solutions. Note that $f(1) = 1$ and so $f(n) \geq 1$ for all n . Consider $g(n) = \log f(n)$, so that $g(n+1) \geq g(n)$ for all n , $g(mn) = g(m) + g(n)$ and $g(n) \geq 0$ for all n . Fix different primes p, q and consider arbitrary positive integers a, b . If $p^a \leq q^b$, then $g(p^a) \leq g(q^b)$, which becomes $ag(p) \leq bg(q)$, or equivalently $a \leq b \frac{g(q)}{g(p)}$. Thus whenever $x = \frac{a}{b}$ is a positive rational number such that $x \leq \frac{\log q}{\log p}$ we also have $x \leq \frac{g(q)}{g(p)}$. Since the number $\frac{\log q}{\log p}$ can be approximated at any order by rational numbers, we conclude that

$$\frac{\log q}{\log p} \leq \frac{g(q)}{g(p)}.$$

Arguing similarly (using a, b such that $p^a \geq q^b$) yields the opposite inequality

$$\frac{\log q}{\log p} \geq \frac{g(q)}{g(p)}, \quad \text{so} \quad \frac{\log q}{\log p} = \frac{g(q)}{g(p)}.$$

We deduce that $\frac{g(p)}{\log p}$ is independent of the choice of the prime p , say equal to some $k \geq 0$ for all p . Then $g(p) = p^k$ for all p and since g is totally multiplicative we conclude that $g(n) = n^k$ for all n , as desired. \square

44. (Erdős) Let $f : \mathbf{N} \rightarrow \mathbf{R}$ be a nonzero multiplicative function such that $f(n+1) \geq f(n)$ for all n . Then there is a nonnegative real number k such that $f(n) = n^k$ for all n .

Proof. Since f is multiplicative and nonzero, we have $f(1) = 1$, and using the hypothesis of the problem we obtain $f(n) \geq 1$ for all $n \geq 1$. We will prove that f is totally multiplicative, which will be enough to conclude thanks to the previous example. For this, we will prove that for any prime p and any $k \geq 1$ we have

$$f(p^{k+1}) = f(p)f(p^k).$$

Fix such p and $k \geq 1$. For any integer $n \geq 1$ not divisible by p we have

$$\begin{aligned} f(n+p)f(p^k)f(p) &= f(np^k + p^{k+1})f(p) \geq f(p^kn + 1)f(p) \\ &= f(p^{k+1}n + p) \geq f(p^{k+1}n) \\ &= f(p^{k+1})f(n). \end{aligned}$$

Similarly,

$$\begin{aligned} f(p^{k+1})f(n+p) &= f(p^{k+1}n + p^{k+2}) \geq f(p^{k+1}n + p) \\ &= f(p)f(p^kn + 1) \geq f(p)f(p^kn) \\ &= f(p)f(p^k)f(n), \end{aligned}$$

We deduce that setting

$$a = \frac{f(p^{k+1})}{f(p)f(p^k)}, \quad b = \frac{1}{a} = \frac{f(p)f(p^k)}{f(p^{k+1})},$$

we have

$$f(n+p) \geq af(n), \quad f(n+p) \geq bf(n)$$

for all n relatively prime to p . Iterating the first inequality yields

$$f(n+jp) \geq a^j f(n)$$

for all $j \geq 1$ and all n relatively prime to p . Taking $j = \left\lfloor \frac{n}{p} \right\rfloor$ we have $f(n+jp) \leq f(2n)$ and so

$$f(2n) \geq a^{\left\lfloor \frac{n}{p} \right\rfloor} f(n).$$

Choosing n odd, the previous inequality becomes $a^{\lfloor \frac{n}{p} \rfloor} \leq f(2)$. Choosing n very large (relatively prime to $2p$), we deduce that $a \leq 1$. Similarly, we obtain $b \leq 1$, which yields $a = b = 1$ and so

$$f(p^{k+1}) = f(p)f(p^k).$$

Since p was an arbitrary prime and k an arbitrary positive integer, we deduce that f is totally multiplicative, as desired. \square

45. Are there infinitely many $n > 1$ such that $n \mid 2^{\sigma(n)} - 1$?

Proof. Let F_i be the i th Fermat number and choose arbitrary prime factors q_0, q_1, \dots of F_0, F_1, \dots , so for instance $q_0 = 3$, $q_1 = 5$, etc. Define $n_d = q_0 q_1 \dots q_d$ for all $d \geq 1$. We claim that $n_d \mid 2^{\sigma(n_d)} - 1$. Since $\sigma(n_d)$ is a multiple of 2^d , it suffices to prove that $q_0 q_1 \dots q_d \mid 2^{2^d} - 1$. Since the Fermat numbers are pairwise relatively prime, so are q_0, \dots, q_d , thus it suffices to prove that each of the numbers q_0, \dots, q_d divides $2^{2^d} - 1$. This is clear, since $2^{2^d} - 1$ is a multiple of F_{i-1} for $i \leq d$. \square

46. An integer $n > 1$ is called perfect if $\sigma(n) = 2n$. Prove that an even number $n > 1$ is perfect if and only if $n = 2^{p-1}(2^p - 1)$, with $2^p - 1$ prime.

Proof. Suppose first that $n = 2^{p-1}(2^p - 1)$, with $2^p - 1$ prime. Since σ is multiplicative, we have

$$\sigma(n) = \sigma(2^{p-1}) \cdot \sigma(2^p - 1) = \frac{2^p - 1}{2 - 1} \cdot 2^p = 2n,$$

hence n is perfect. The converse is more difficult. Suppose that $n = 2^k m$ is perfect, with $k \geq 1$ and m odd. Again, by multiplicativity of σ we have

$$2^{k+1}m = 2n = \sigma(2^k)\sigma(m) = (2^{k+1} - 1)\sigma(m).$$

Since $\gcd(2^{k+1}, 2^{k+1} - 1) = 1$, there is an integer a such that $m = a(2^{k+1} - 1)$ and $\sigma(m) = 2^{k+1}a$. If $a > 1$, then $1, a$ and m are divisors of

m , hence $\sigma(m) \geq 1 + a + m = 1 + 2^{k+1}a$, a contradiction. Hence $a = 1$, $m = 2^{k+1} - 1$ and $\sigma(m) = 2^{k+1} = m + 1$. The last equality implies that m is a prime, which finishes the proof, since $n = 2^k m = 2^k(2^{k+1} - 1)$. \square

47. Let n be an even positive integer. Prove that $\sigma(\sigma(n)) = 2n$ if and only if there is a prime p such that $2^p - 1$ is a prime and $n = 2^{p-1}$.

Proof. Suppose that $n = 2^{p-1}$ with $2^p - 1$ prime. Then $\sigma(n) = 2^p - 1$ and $\sigma(\sigma(n)) = 1 + 2^p - 1 = 2^p = 2n$. Conversely, suppose that $\sigma(\sigma(n)) = 2n$ and write $n = 2^k m$, with $k \geq 1$ and m odd. Suppose by contradiction that $m > 1$ and note that the condition $\sigma(\sigma(n)) = 2n$ can be written

$$\sigma((2^{k+1} - 1)\sigma(m)) = 2^{k+1}m.$$

Since $1, \sigma(m)$ and $(2^{k+1} - 1)\sigma(m)$ are different divisors of $(2^{k+1} - 1)\sigma(m)$, we deduce that

$$2^{k+1}m \geq 1 + \sigma(m) + (2^{k+1} - 1)\sigma(m) > 2^{k+1}\sigma(m) > 2^{k+1}m,$$

a contradiction. Hence $m = 1$ and $n = 2^{k+1}$, with $\sigma(2^{k+1} - 1) = 2^{k+1}$. This clearly implies that $2^{k+1} - 1$ is a prime, hence $k + 1 = p$ is a prime. The result follows. \square

48. (Romania TST 2010) Prove that for each positive integer a we have $\sigma(an) < \sigma(an + 1)$ for infinitely many positive integers n .

Proof. The idea is to choose n prime (so that an has few divisors) such that $an + 1$ has many prime divisors. Suppose that p_1, \dots, p_k are pairwise distinct primes that do not divide a and that $n > a$ is a prime such that $an + 1 \equiv 0 \pmod{p_1 \dots p_k}$. Then

$$\sigma(an + 1) \geq (an + 1) \cdot \prod_{i=1}^k \left(1 + \frac{1}{p_i}\right) > an \cdot \prod_{i=1}^k \left(1 + \frac{1}{p_i}\right).$$

On the other hand

$$\sigma(an) = \sigma(a)\sigma(n) = \sigma(a)(1 + n) < 2\sigma(a)n.$$

It is thus enough to ensure that

$$\prod_{i=1}^k \left(1 + \frac{1}{p_i}\right) > \frac{2\sigma(a)}{a}$$

in order to have $\sigma(an) < \sigma(an+1)$. It is now clear how to proceed: let p_1, p_2, \dots be the increasing sequence of primes that do not divide a . Since only finitely many primes divide a , by theorem 4.74 there is k such that

$$\prod_{i=1}^k \left(1 + \frac{1}{p_i}\right) > \sum_{i=1}^k \frac{1}{p_i} > \frac{2\sigma(a)}{a}.$$

Fixing such a k , Dirichlet's theorem yields the existence of infinitely many primes n such that $an+1 \equiv 0 \pmod{p_1 p_2 \dots p_k}$. The result follows. \square

49. (IMO Shortlist 2004) Prove that for infinitely many positive integers a the equation $\tau(an) = n$ has no solutions in positive integers.

Proof. We will prove that if $a = p^{p-1}$, with $p > 3$, then the equation has no solutions. Assume that n is a solution and let $m = an$, so that $a\tau(m) = m$. Since a divides m , we can write $m = p^r s$ with $r \geq p-1$ and s relatively prime to p . Then the equation becomes

$$(r+1)\tau(s) = p^{r-p+1}s.$$

This forces $r \geq p$ (otherwise $r = p-1$ and the right-hand side is not a multiple of p , while the left-hand side is divisible by p). Let $k = r - p + 1$, so that $k \geq 1$ and

$$(k+p)\tau(s) = p^k s.$$

Since $\tau(s) \leq s$, we deduce that $k+p \geq p^k$. Assume that $k \geq 2$, then

$$p^k - p = p(p^{k-1} - 1) \geq 3(3^{k-1} - 1) \geq 3 \cdot 2(k-1) > k,$$

a contradiction. Thus $k = 1$ and $(p+1)\tau(s) = ps$. Write now the prime factorization of s ,

$$s = p_1^{a_1} \dots p_d^{a_d}.$$

Then for all i we have

$$\frac{p}{p+1} = \frac{\tau(s)}{s} = \prod_j \frac{a_j + 1}{p_j^{a_j}} \leq \frac{a_i + 1}{p_i^{a_i}} \leq \frac{a_i + 1}{2^{a_i}}.$$

On the other hand

$$2^{a_i} \geq 1 + a_i + \binom{a_i}{2}.$$

Combining these inequalities yields

$$p \binom{a_i}{2} \leq a_i + 1$$

for all i . Since $p \geq 4$, this immediately implies $a_i = 1$ for all i . But then the equation $(p+1)\tau(s) = ps$ becomes

$$(p+1)2^d = p \cdot p_1 \cdots p_d.$$

We deduce that $p \mid 2^d(p+1)$, which is obviously impossible. Therefore for such a the equation has no solution and the result follows. \square

50. (IMO) Let $\tau(n)$ be the number of divisors of a positive integer n . Find all positive integers k such that $k = \frac{\tau(n^2)}{\tau(n)}$ for some n .

Proof. Answer: all odd positive integers k . Let $k = \frac{\tau(n^2)}{\tau(n)}$ for some n . If $n = 1$ then $k = 1$. If $n > 1$ and $n = p_1^{r_1} \cdots p_s^{r_s}$ is the prime decomposition of n then $\tau(n^2) = (2r_1 + 1) \cdots (2r_s + 1)$ is an odd number and hence k is odd. Conversely, let $k = 2m + 1$ is an odd number. We shall prove by induction on m that there are r_1, \dots, r_s and hence n such that

$$k = \frac{(2r_1 + 1) \cdots (2r_s + 1)}{(r_1 + 1) \cdots (r_s + 1)} = \frac{\tau(n^2)}{\tau(n)}.$$

If $m = 1$ then

$$3 = \frac{(2 \cdot 2 + 1)(2 \cdot 4 + 1)}{(2 + 1)(4 + 1)}.$$

Suppose that for all $m < M$ we can write $2m + 1$ as a fraction of the desired form and let $k = 2M + 1$. If $k + 1 = 2^l \cdot t$, where t is odd then

$$t = \frac{k+1}{2^l} \leq \frac{k+1}{2} < k.$$

Consider the numbers

$$r_1 = 2^l - 2^0 t - 2^0, r_2 = 2r_1, \dots, r_l = 2^{l-1} r_1.$$

Then for $n_1 = p_1^{r_1} \dots p_l^{r_l}$ we have

$$k_1 = \frac{\tau(n_1^2)}{\tau(n_1)} = \frac{(2r_1 + 1) \dots (2r_l + 1)}{(r_1 + 1) \dots (r_l + 1)} = \frac{2r_l + 1}{r_1 + 1} = \frac{2^l t - 1}{t}.$$

Since $t < k$ we now that there is $n_2 = q_1^{\alpha_1} \dots q_s^{\alpha_s}$ such that $t = \frac{\tau(n_2^2)}{\tau(n_2)}$. Then choosing the primes p_1, \dots, p_l different from q_1, \dots, q_s we set $n = n_1 n_2$ and get

$$\frac{\tau(n^2)}{\tau(n)} = \frac{\tau(n_1^2)}{\tau(n_1)} \cdot \frac{\tau(n_2^2)}{\tau(n_2)} = k_1 t = 2^l t - 1 = k.$$

Hence the induction is finished and the statement is proved. \square

51. A positive integer a is called highly divisible if it has more divisors than any number less than a . If p is a prime number and $a > 1$ is an integer, we write $v_p(a)$ for the exponent of p in the prime factorization of a . Prove that
- There are infinitely many highly divisible numbers.
 - If a is highly divisible and $p < q$ are primes, then $v_p(a) \geq v_q(a)$.
 - Let p, q be primes such that $p^k < q$ for some positive integer k . Prove that if a is highly divisible and a multiple of q , then a is a multiple of p^k .
 - Let p, q be primes and let k be a positive integer such that $p^k > q$. Prove that if p^{2k} divides some highly divisible number a , then q divides a .
 - (China TST 2012) Let n be a positive integer. Prove that all sufficiently large highly divisible numbers are multiples of n .

Proof. We will constantly use the formula

$$\tau(x) = \prod_{p|x} (1 + v_p(x))$$

for the number of divisors $\tau(x)$ of x .

a) Suppose that there is a largest highly divisible number a . Then for $b > a$ we have $\tau(b) \leq \max_{j < b} \tau(j)$, hence the sequence $(\tau(b))_{b > a}$ is bounded. This is clearly absurd.

b) If $v_p(a) < v_q(a)$, then $b = \frac{a}{p^{v_p(a)} q^{v_q(a)}} p^{v_q(a)} q^{v_p(a)}$ is less than a and $\tau(b) = \tau(a)$, contradicting the fact that a is highly divisible.

c) Let $b = \frac{ap^k}{q}$. Note that $b < a$, hence $\tau(b) < \tau(a)$, since a is highly divisible. We deduce that

$$v_q(a)(v_p(a) + k + 1) < (1 + v_p(a))(1 + v_q(a)),$$

which simplifies to $kv_q(a) < 1 + v_p(a)$. Since $v_q(a) \geq 1$ by assumption, it follows that $v_p(a) \geq k$, thus p^k divides a and we are done.

d) Suppose that q does not divide a and let $b = \frac{aq}{p^k}$. Again, $b < a$ hence $\tau(b) < \tau(a)$, which translates into

$$(1 + v_p(a) - k)(2 + v_q(a)) < 1 + v_p(a).$$

Since $v_q(a) = 0$, this reduces to $1 + v_p(a) < 2k$, contradicting the fact that p^{2k} divides a .

e) Let p_1, p_2, \dots be the increasing sequence of primes. It suffices to prove that for all n and k , all sufficiently large highly divisible numbers are multiples of $(p_1 \dots p_n)^k$. By part b), it suffices to ensure that such numbers are multiples of p_n^k . Suppose that this is not the case, hence infinitely many highly divisible numbers a_i are not multiples of p_n^k . Let q be a prime greater than p_n^k . By part c), a_i are not multiples of q , hence their prime factors are all less than q by part a). Let q_1, \dots, q_s be all primes less than q and let m be such that $q_1^m > q$. If a_i is sufficiently large, then at least one of the numbers $v_{q_1}(a_i), \dots, v_{q_s}(a_i)$ is greater than $2m$. By part d) it follows that q divides a_i , a contradiction. The result follows. \square

52. Let $n > 1$ be an integer. Compute

$$\sum_{d|n} (-1)^{\frac{n}{d}} \varphi(d).$$

Proof. If n is odd, then so is $\frac{n}{d}$ for all $d | n$, hence

$$\sum_{d|n} (-1)^{\frac{n}{d}} \varphi(d) = - \sum_{d|n} \varphi(d) = -n$$

by Gauss' theorem 4.112. Suppose that n is even and write $n = 2^k m$ with $k \geq 1$ and m odd. Then $\frac{n}{d}$ is odd if and only if $v_2(d) = k$, that is $d = 2^k e$ with $e | m$. Hence

$$\sum_{d|n} (-1)^{\frac{n}{d}} \varphi(d) = \sum_{\substack{d|n \\ v_2(d) < k}} \varphi(d) - \sum_{e|m} \varphi(2^k e) = \sum_{d|n} \varphi(d) - 2 \sum_{e|m} 2^{k-1} \varphi(e).$$

Using Gauss' theorem 4.112 twice, we obtain

$$\sum_{d|n} (-1)^{\frac{n}{d}} \varphi(d) = n - 2^k m = 0. \quad \square$$

53. (IMO 1991) Let $1 = a_1 < a_2 < \dots < a_{\varphi(n)}$ be the totatives of $n > 1$. Prove that $a_1, a_2, \dots, a_{\varphi(n)}$ form an arithmetic progression if and only if n is either 6, a prime number or a power of 2.

Proof. It is clear that if $n = 6$, a prime or a power of 2, then $a_1, \dots, a_{\varphi(n)}$ form an arithmetic progression, so let us prove the converse. The case $n \leq 6$ being easy, we assume that $n \geq 7$. Then $\varphi(n) \geq 3$. If $a_2 = 2$, then $a_1, a_2, \dots, a_{\varphi(n)}$ must be consecutive numbers and so, since $a_{\varphi(n)} = n - 1$, we must have $\varphi(n) = n - 1$. Thus n is relatively prime to $1, 2, \dots, n - 1$, and so n is a prime. If $a_2 = 3$, then we must have $a_j = 2j - 1$ for all j , so $n - 1 = 2\varphi(n) - 1$ and $n = 2\varphi(n)$. Write $n = 2^k m$ with $k \geq 1$ and m odd, then the equation becomes $m = \varphi(m)$, with the unique solution $m = 1$, so n is a power of 2.

Assume now that $a_2 > 3$. Thus n is a multiple of 3. Moreover,

$$n - 1 = a_{\varphi(n)} = 1 + (\varphi(n) - 1)(a_2 - 1).$$

Note that 3 does not divide a_2 , and the last relation shows that 3 does not divide $a_2 - 1$, thus $a_2 \equiv 2 \pmod{3}$. But then $a_3 = 2a_2 - 1 \equiv 0 \pmod{3}$, contradicting the fact that $\gcd(a_3, n) = 1$. Thus this case does not lead to any solution, and the result follows. \square

54. Let $n \geq 2$. Prove that n is a prime if and only if $\varphi(n) \mid n - 1$ and $n + 1 \mid \sigma(n)$ (recall that $\sigma(n)$ is the sum of the positive divisors of n).

Proof. One direction being obvious, suppose that $\varphi(n) \mid n - 1$ and $n + 1 \mid \sigma(n)$ and let p_1, \dots, p_k be the (pairwise distinct) prime factors of n . If there is i such that $p_i^2 \mid n$, then $p_i \mid \varphi(n) \mid n - 1$, a contradiction. Hence $n = p_1 p_2 \dots p_k$. Suppose that $k > 1$. Note that n is odd, since $\varphi(n)$ is even (the case $n = 2$ is excluded by the hypothesis $k > 1$) and $\varphi(n) \mid n - 1$. Thus all p_i are odd, hence $2^k \mid \varphi(n) = (p_1 - 1) \dots (p_k - 1)$. In particular $4 \mid \varphi(n) \mid n - 1$ and so 4 does not divide $n + 1$. Also, $2^k \mid \sigma(n) = (p_1 + 1) \dots (p_k + 1)$. Combining the last two observations, we deduce that 2^{k-1} divides $\frac{\sigma(n)}{n+1}$, hence $\frac{\sigma(n)}{n+1} \geq 2^{k-1}$. This is however absurd, since

$$\frac{\sigma(n)}{n+1} < \frac{\sigma(n)}{n} = \left(1 + \frac{1}{p_1}\right) \cdot \dots \cdot \left(1 + \frac{1}{p_k}\right) \leq \left(\frac{4}{3}\right)^k < 2^{k-1}. \quad \square$$

Remark 8.13. A famous conjecture of Lehmer asserts that an integer $n > 1$ is a prime if and only if $\varphi(n)$ divides $n - 1$ (of course, only one implication is difficult). This is still open, even though one knows that the possible counterexamples are huge.

55. Let k be a positive integer. Prove that there is a positive integer n such that $\varphi(n) = \varphi(n + k)$.

Proof. Let p be the smallest prime not dividing k and choose $n = (p-1)k$. Then $\varphi(n + k) = \varphi(pk) = (p-1)\varphi(k)$. On the other hand,

$$\varphi(n) = \varphi((p-1)k) = (p-1)k \prod_{q \mid (p-1)k} \left(1 - \frac{1}{q}\right)$$

and since all prime factors q of $p-1$ are prime divisors of k (by minimality of p) we deduce that

$$\varphi(n) = (p-1)k \prod_{q|k} \left(1 - \frac{1}{q}\right) = (p-1)\varphi(k) = \varphi(n+k),$$

solving the problem. □

56. Prove that for all $n \geq 1$ we have

$$\frac{\varphi(1)}{2^1 - 1} + \frac{\varphi(2)}{2^2 - 1} + \dots + \frac{\varphi(n)}{2^n - 1} < 2.$$

Proof. The key observation is that

$$\frac{\varphi(k)}{2^k - 1} = \frac{\varphi(k)}{2^k} \cdot \frac{1}{1 - \frac{1}{2^k}} = \varphi(k) \sum_{j \geq 1} \frac{1}{2^{jk}}.$$

Hence

$$\sum_{k=1}^n \frac{\varphi(k)}{2^k - 1} = \sum_{k=1}^n \sum_{j \geq 1} \frac{\varphi(k)}{2^{jk}} = \sum_{d \geq 1} \frac{1}{2^d} \sum_{jk=d, k \leq n} \varphi(k).$$

Now for all $d \geq 1$ we clearly have (using Gauss' theorem 4.112)

$$\sum_{jk=d, k \leq n} \varphi(k) \leq \sum_{k|d} \varphi(k) = d,$$

thus

$$\sum_{k=1}^n \frac{\varphi(k)}{2^k - 1} \leq \sum_{d \geq 1} \frac{d}{2^d}.$$

Since

$$x + 2x^2 + \dots + nx^n = \frac{nx^{n+2} - (n+1)x^{n+1} + x}{(x-1)^2},$$

we deduce by choosing $x = 1/2$ and letting $n \rightarrow \infty$ that

$$\sum_{d=1}^{\infty} \frac{d}{2^d} = 2,$$

and the result follows. □

Remark 8.14. The argument shows that

$$\sum_{n \geq 1} \frac{\varphi(n)}{2^n - 1} = 2,$$

and, more generally, that for all $x \in (-1, 1)$ we have

$$\sum_{n \geq 1} \varphi(n) \frac{x^n}{1 - x^n} = \frac{x}{(1 - x)^2}.$$

57. a) Prove that there are infinitely many integers $n > 1$ such that

$$\varphi(n) \geq \varphi(k) + \varphi(n - k)$$

for all $1 \leq k \leq n - 1$.

b) Are there infinitely many $n > 1$ such that $\varphi(n) \leq \varphi(k) + \varphi(n - k)$ for all $1 \leq k \leq n - 1$?

Proof. a) We claim that any odd prime p has this property. Indeed, we have

$$\varphi(k) + \varphi(p - k) \leq k - 1 + p - k - 1 = p - 2 < \varphi(p) = p - 1.$$

b) The answer is positive. Let p_1, p_2, \dots be the increasing sequence of primes and define $n_d = p_1 p_2 \dots p_d$. We will prove that this is a solution of the problem for $d \geq 2$. Choose any $k \in \{1, 2, \dots, n_d - 1\}$ and let $q_1 < q_2 < \dots < q_l$ be the prime factors of k . Note that $q_1 \geq p_1, q_2 \geq p_2, \dots$ and since

$$p_1 \dots p_d = n_d > k \geq q_1 \dots q_l \geq p_1 p_2 \dots p_l$$

we must have $l < d$. We deduce that

$$\frac{\varphi(k)}{k} = \prod_{i=1}^l \left(1 - \frac{1}{q_i}\right) \geq \prod_{i=1}^d \left(1 - \frac{1}{p_i}\right) = \frac{\varphi(n_d)}{n_d}.$$

Since a similar inequality holds with $n_d - k$ instead of k , we conclude that

$$\varphi(k) + \varphi(n_d - k) \geq k \cdot \frac{\varphi(n_d)}{n_d} + (n_d - k) \cdot \frac{\varphi(n_d)}{n_d} = \varphi(n_d),$$

proving that n_d is a solution of the problem. \square

58. (AMM 11544) Prove that for any integer $m > 1$ we have

$$\sum_{k=0}^{m-1} \varphi(2k+1) \left\lfloor \frac{m+k}{2k+1} \right\rfloor = m^2.$$

Proof. Denote by x_m the left-hand side of the equality. Then

$$x_{m+1} - x_m = \varphi(2m+1) + \sum_{k=0}^{m-1} \varphi(2k+1) \left(\left\lfloor \frac{m+k+1}{2k+1} \right\rfloor - \left\lfloor \frac{m+k}{2k+1} \right\rfloor \right).$$

Recall that in general $\left\lfloor \frac{n+1}{k} \right\rfloor - \left\lfloor \frac{n}{k} \right\rfloor = 1$ if and only if $k \mid n+1$, thus

$$x_{m+1} - x_m = \varphi(2m+1) + \sum_{\substack{0 \leq k \leq m-1 \\ 2k+1 \mid m+k+1}} \varphi(2k+1).$$

The condition $2k+1 \mid m+k+1$ is equivalent to $2k+1 \mid 2(m+k+1)$ and also with $2k+1 \mid 2m+1$. Since all positive divisors of $2k+1$ are odd, we obtain

$$\varphi(2m+1) + \sum_{\substack{0 \leq k \leq m-1 \\ 2k+1 \mid m+k+1}} \varphi(2k+1) = \sum_{d \mid 2m+1} \varphi(d) = 2m+1,$$

the last equality by Gauss' theorem 4.112. Thus

$$x_{m+1} - x_m = 2m+1$$

and the result follows by induction. \square

Remark 8.14. The argument shows that

$$\sum_{n \geq 1} \frac{\varphi(n)}{2^n - 1} = 2,$$

and, more generally, that for all $x \in (-1, 1)$ we have

$$\sum_{n \geq 1} \varphi(n) \frac{x^n}{1 - x^n} = \frac{x}{(1 - x)^2}.$$

57. a) Prove that there are infinitely many integers $n > 1$ such that

$$\varphi(n) \geq \varphi(k) + \varphi(n - k)$$

for all $1 \leq k \leq n - 1$.

b) Are there infinitely many $n > 1$ such that $\varphi(n) \leq \varphi(k) + \varphi(n - k)$ for all $1 \leq k \leq n - 1$?

Proof. a) We claim that any odd prime p has this property. Indeed, we have

$$\varphi(k) + \varphi(p - k) \leq k - 1 + p - k - 1 = p - 2 < \varphi(p) = p - 1.$$

b) The answer is positive. Let p_1, p_2, \dots be the increasing sequence of primes and define $n_d = p_1 p_2 \dots p_d$. We will prove that this is a solution of the problem for $d \geq 2$. Choose any $k \in \{1, 2, \dots, n_d - 1\}$ and let $q_1 < q_2 < \dots < q_l$ be the prime factors of k . Note that $q_1 \geq p_1, q_2 \geq p_2, \dots$ and since

$$p_1 \dots p_d = n_d > k \geq q_1 \dots q_l \geq p_1 p_2 \dots p_l$$

we must have $l < d$. We deduce that

$$\frac{\varphi(k)}{k} = \prod_{i=1}^l \left(1 - \frac{1}{q_i}\right) \geq \prod_{i=1}^d \left(1 - \frac{1}{p_i}\right) = \frac{\varphi(n_d)}{n_d}.$$

Since a similar inequality holds with $n_d - k$ instead of k , we conclude that

$$\varphi(k) + \varphi(n_d - k) \geq k \cdot \frac{\varphi(n_d)}{n_d} + (n_d - k) \cdot \frac{\varphi(n_d)}{n_d} = \varphi(n_d),$$

proving that n_d is a solution of the problem. \square

58. (AMM 11544) Prove that for any integer $m > 1$ we have

$$\sum_{k=0}^{m-1} \varphi(2k+1) \left\lfloor \frac{m+k}{2k+1} \right\rfloor = m^2.$$

Proof. Denote by x_m the left-hand side of the equality. Then

$$x_{m+1} - x_m = \varphi(2m+1) + \sum_{k=0}^{m-1} \varphi(2k+1) \left(\left\lfloor \frac{m+k+1}{2k+1} \right\rfloor - \left\lfloor \frac{m+k}{2k+1} \right\rfloor \right).$$

Recall that in general $\left\lfloor \frac{n+1}{k} \right\rfloor - \left\lfloor \frac{n}{k} \right\rfloor = 1$ if and only if $k \mid n+1$, thus

$$x_{m+1} - x_m = \varphi(2m+1) + \sum_{\substack{0 \leq k \leq m-1 \\ 2k+1 \mid m+k+1}} \varphi(2k+1).$$

The condition $2k+1 \mid m+k+1$ is equivalent to $2k+1 \mid 2(m+k+1)$ and also with $2k+1 \mid 2m+1$. Since all positive divisors of $2k+1$ are odd, we obtain

$$\varphi(2m+1) + \sum_{\substack{0 \leq k \leq m-1 \\ 2k+1 \mid m+k+1}} \varphi(2k+1) = \sum_{d \mid 2m+1} \varphi(d) = 2m+1,$$

the last equality by Gauss' theorem 4.112. Thus

$$x_{m+1} - x_m = 2m+1$$

and the result follows by induction. \square

59. a) Prove that for all $n > 1$ we have

$$2 \sum_{k=1}^n \varphi(k) = 1 + \sum_{k=1}^n \mu(k) \left\lfloor \frac{n}{k} \right\rfloor^2.$$

b) Prove that for all $n > 1$ we have

$$\left| \varphi(1) + \varphi(2) + \dots + \varphi(n) - \frac{3}{\pi^2} n^2 \right| < 2n + n \log n.$$

Proof. a) The identity

$$\varphi(k) = k \cdot \sum_{d|k} \frac{\mu(d)}{d}$$

gives

$$\begin{aligned} \sum_{k=1}^n \varphi(k) &= \sum_{k=1}^n k \cdot \sum_{d|k} \frac{\mu(d)}{d} = \sum_{d=1}^n \frac{\mu(d)}{d} \sum_{\substack{k \leq n \\ d|k}} k \\ &= \sum_{d=1}^n \frac{\mu(d)}{d} \sum_{j \leq \lfloor \frac{n}{d} \rfloor} (jd) = \sum_{d=1}^n \mu(d) \frac{\lfloor \frac{n}{d} \rfloor (\lfloor \frac{n}{d} \rfloor + 1)}{2}. \end{aligned}$$

Thus it remains to prove that

$$\sum_{d=1}^n \mu(d) \left\lfloor \frac{n}{d} \right\rfloor = 1$$

for $n > 1$. But with a similar argument we obtain

$$\sum_{d=1}^n \mu(d) \left\lfloor \frac{n}{d} \right\rfloor = \sum_{d=1}^n \mu(d) \sum_{\substack{k \leq n \\ d|k}} 1 = \sum_{k=1}^n \sum_{d|k} \mu(d) = 1$$

since $\sum_{d|k} \mu(d)$ equals 0 for $k > 1$ and 1 for $k = 1$.

b) We will use the inequality

$$\left| \sum_{k \leq n} \frac{\mu(k)}{k^2} - \frac{6}{\pi^2} \right| \leq \frac{1}{n},$$

which can be proved arguing as in the proof of theorem 4.130. Using the inequality $x^2 - [x]^2 \leq 2x$ for $x \geq 0$, part a) and the result taken for granted, we obtain

$$\begin{aligned} \left| \varphi(1) + \varphi(2) + \dots + \varphi(n) - \frac{3}{\pi^2} n^2 \right| &= \frac{1}{2} \left| 1 + \sum_{k=1}^n \mu(k) \left[\frac{n}{k} \right]^2 - \frac{6}{\pi^2} n^2 \right| \\ &< \frac{n+1}{2} + \sum_{k=1}^n \frac{1}{2} \left(\frac{n^2}{k^2} - \left[\frac{n}{k} \right]^2 \right) \leq \frac{n+1}{2} + \sum_{k=1}^n \frac{n}{k} < 2n + n \log n, \end{aligned}$$

using the inequality

$$\sum_{k=1}^n \frac{1}{k} \leq 1 + \log n. \quad \square$$

60. Let $a_1, \dots, a_{\varphi(n)}$ be the totatives of $n > 1$.

a) Prove that for all $m \geq 1$ we have

$$a_1^m + a_2^m + \dots + a_{\varphi(n)}^m = \sum_{d|n} \mu(d) d^m \left(1^m + 2^m + \dots + \left(\frac{n}{d} \right)^m \right).$$

b) Compute $a_1^2 + a_2^2 + \dots + a_{\varphi(n)}^2$.

Proof. a) Note that

$$\sum_{k=1}^n k^m = \sum_{d|n} \sum_{\substack{1 \leq k \leq n \\ \gcd(k,n)=d}} k^m = \sum_{d|n} d^m \sum_{\substack{1 \leq j \leq \frac{n}{d} \\ \gcd(j, \frac{n}{d})=1}} j^m.$$

In other words, if

$$S_m(n) = \sum_{\substack{1 \leq k \leq n \\ \gcd(k,n)=1}} k^m,$$

then

$$\frac{1}{n^m} \sum_{k=1}^n k^m = \sum_{d|n} \frac{S_m\left(\frac{n}{d}\right)}{\left(\frac{n}{d}\right)^m} = \sum_{d|n} \frac{S_m(d)}{d^m}.$$

The desired identity then follows from Möbius' inversion formula.

b) Taking $m = 2$ in part a) yields

$$\begin{aligned} a_1^2 + a_2^2 + \dots + a_{\varphi(n)}^2 &= \sum_{d|n} \mu(d) d^2 \left(1^2 + \dots + \frac{n^2}{d^2} \right) \\ &= \sum_{d|n} \mu(d) d^2 \cdot \frac{\frac{n}{d} \left(\frac{n}{d} + 1 \right) \left(\frac{2n}{d} + 1 \right)}{6} \\ &= \frac{n^3}{3} \sum_{d|n} \frac{\mu(d)}{d} + \frac{n^2}{2} \sum_{d|n} \mu(d) + \frac{n}{6} \sum_{d|n} d \mu(d). \end{aligned}$$

Using the identities (the second one uses the hypothesis $n > 1$)

$$\sum_{d|n} \frac{\mu(d)}{d} = \frac{\varphi(n)}{n}, \quad \sum_{d|n} \mu(d) = 0$$

and

$$\sum_{d|n} d \mu(d) = \prod_{p|n} (1 - p),$$

we conclude that

$$a_1^2 + a_2^2 + \dots + a_{\varphi(n)}^2 = \frac{n^2 \varphi(n)}{3} + \frac{n}{6} \prod_{p|n} (1 - p). \quad \square$$

Remark 8.15. The identity obtained in part a) can be used to prove the equidistribution of the numbers $\frac{a_1}{n}, \frac{a_2}{n}, \dots, \frac{a_{\varphi(n)}}{n}$ as $n \rightarrow \infty$. This means that for each interval $I \subset [0, 1]$ we have

$$\lim_{n \rightarrow \infty} \frac{|I \cap \{\frac{a_1}{n}, \frac{a_2}{n}, \dots, \frac{a_{\varphi(n)}}{n}\}|}{\varphi(n)} = \text{length}(I)$$

or equivalently

$$\lim_{n \rightarrow \infty} \frac{1}{\varphi(n)} \sum_{i=1}^{\varphi(n)} f\left(\frac{a_i}{n}\right) = \int_0^1 f(x) dx$$

for all continuous functions $f : [0, 1] \rightarrow \mathbf{R}$. Finally, this reduces (thanks to the Weierstrass approximation theorem) to proving that

$$\lim_{n \rightarrow \infty} \frac{1}{\varphi(n)} \sum_{i=1}^{\varphi(n)} \left(\frac{a_i}{n}\right)^m = \frac{1}{m+1}$$

for all $m \geq 0$, which is a not too difficult (but technical enough) consequence of part a) of the previous exercise.

61. (Serbia 2011) Prove that if $n > 1$ is odd and $\varphi(n)$, $\varphi(n+1)$ are powers of 2, then $n+1$ is a power of 2 or $n=5$.

Proof. Observe first that if a is an odd integer such that $\varphi(a)$ is a power of 2, then a is squarefree and all of its prime factors are Fermat numbers. Indeed, if $a = p_1^{k_1} \dots p_r^{k_r}$ is the prime factorization of a , then $p_1^{k_1-1} \dots p_r^{k_r-1} (p_1-1) \dots (p_r-1)$ being a power of 2 forces $k_1 = \dots = k_r = 1$ and all $p_i - 1$ being powers of 2, i.e. p_i are Fermat primes. Thus we can write $a = p_1 \dots p_r$ and assuming that $p_1 < \dots < p_r$, we obtain $p_{i+1} - 1 \geq (p_i - 1)^2$ (since $p_i = 2^{2^{a_i}} + 1$ for some $a_1 < \dots < a_r$). We deduce that

$$\begin{aligned} \frac{p_1}{p_1-1} &\leq \frac{a}{\varphi(a)} = \prod_{i=1}^r \frac{p_i}{p_i-1} \\ &\leq \left(1 + \frac{1}{p_1-1}\right) \cdot \left(1 + \frac{1}{(p_1-1)^2}\right) \cdot \left(1 + \frac{1}{(p_1-1)^4}\right) \cdot \dots = \frac{p_1-1}{p_1-2}, \end{aligned}$$

the last equality being a consequence of the general formula (valid for $x \in (0, 1)$)

$$(1+x)(1+x^2)(1+x^4)\dots = \frac{1}{1-x},$$

which follows directly from the difference of squares formula. We conclude that if a is an odd integer for which $\varphi(a)$ is a power of 2, then

$$\varphi(a) \frac{p_1}{p_1 - 1} \leq a \leq \varphi(a) \cdot \frac{p_1 - 1}{p_1 - 2},$$

where p_1 is the smallest prime factor of a .

Coming back to our problem, write $\varphi(n) = 2^A$ and $\varphi(n+1) = 2^B$, then applying the previous discussion to n and to the largest odd divisor of $n+1$, we obtain

$$2^A < 2^A \frac{p_1}{p_1 - 1} \leq n < 2^A \cdot \frac{p_1 - 1}{p_1 - 2},$$

$$2^{B+1} < 2^{B+1} \frac{q_1}{q_1 - 1} \leq n+1 < 2^{B+1} \frac{q_1 - 1}{q_1 - 2},$$

where p_1 is the smallest prime factor of n and q_1 is the smallest odd prime factor of $n+1$ (assuming that $n+1$ is not a power of 2). Since $p_1, q_1 \geq 3$, we have $\frac{p_1-1}{p_1-2} \leq 2$ and similarly for q_1 . We deduce that

$$2^A < n < 2^{B+2}, \quad 2^{B+1} \leq n < 2^{A+1},$$

which yields $A = B + 1$. Combining this with the previous inequalities yields

$$\frac{q_1 - 1}{q_1 - 2} > \frac{p_1}{p_1 - 1} \quad \text{and} \quad \frac{n+1}{n} > \frac{q_1}{q_1 - 1} \cdot \frac{p_1 - 2}{p_1 - 1}.$$

The first inequality yields $p_1 > q_1 - 1$, while the second one can be rewritten as

$$\frac{1}{n} > \frac{p_1 - q_1 - 1}{(p_1 - 1)(q_1 - 1)}.$$

Since $\gcd(n, n+1) = 1$, we cannot have $p_1 = q_1$, thus necessarily $p_1 \geq q_1 + 2$ (since p_1, q_1 are odd). We conclude that

$$n < (p_1 - 1)(q_1 - 1) < p_1^2.$$

Since p_1 is the smallest prime factor of n , we deduce that $n = p_1$. Thus $p_1 = 1 + 2^A$ and $\varphi(2^A + 2) = 2^B$. Since we assumed that $n+1$ is

not a power of 2, we have $A > 1$, thus $\varphi(2^{A-1} + 1) = 2^B$. But then $2^{A-1} + 1$ (which is odd) is a product of distinct Fermat numbers. Using the uniqueness of the binary expansion, we easily obtain that $2^{A-1} + 1$ must itself be a Fermat prime, i.e. $A - 1$ is a power of 2. But since $1 + 2^A = p_1$ is a prime, A must also be a power of 2. We conclude that $A = 2$ and so $n = 5$. The result follows. \square

62. (Komal A 492) Let A be a finite set of positive integers. Prove that

$$\sum_{S \subset A} (-2)^{|S|-1} \gcd(S) > 0,$$

the sum running over all nonempty subsets S of A and $\gcd(S)$ denoting the greatest common divisor of all elements of S .

Proof. Let $a_1 < \dots < a_n$ be the elements of A and let N be their least common multiple. Let $1_{u|a_i}$ be 1 if $u \mid a_i$ and 0 otherwise. The following relation follows directly from Gauss' formula

$$\gcd(a_{i_1}, \dots, a_{i_k}) = \sum_{u \mid \gcd(a_{i_1}, \dots, a_{i_k})} \varphi(u) = \sum_{u \mid N} \varphi(u) \cdot 1_{u|a_{i_1}} \cdot \dots \cdot 1_{u|a_{i_k}}.$$

We deduce that

$$\begin{aligned} \sum_{S \subset A} (-2)^{|S|-1} \gcd(S) &= \sum_{k=1}^n \sum_{i_1 < \dots < i_k} (-2)^{k-1} \gcd(a_{i_1}, \dots, a_{i_k}) \\ &= \sum_{k=1}^n (-2)^{k-1} \sum_{i_1 < \dots < i_k} \sum_{u \mid N} \varphi(u) \cdot 1_{u|a_{i_1}} \cdot \dots \cdot 1_{u|a_{i_k}} \\ &= \sum_{u \mid N} \varphi(u) \cdot \sum_{k=1}^n (-2)^{k-1} \sum_{i_1 < \dots < i_k} 1_{u|a_{i_1}} \cdot \dots \cdot 1_{u|a_{i_k}} \\ &= \sum_{u \mid N} \varphi(u) \cdot \frac{1 - (1 - 2 \cdot 1_{u|a_1})(1 - 2 \cdot 1_{u|a_2}) \dots (1 - 2 \cdot 1_{u|a_n})}{2}. \end{aligned}$$

All terms in the previous sum are nonnegative, and the term corresponding to $u = a_n$ is equal to $\varphi(a_n)$. We conclude that

$$\sum_{S \subset A} (-2)^{|S|-1} \gcd(S) \geq \varphi(a_n) > 0. \quad \square$$

8.4 Congruences involving prime numbers

1. Prove that for all primes p the number

$$\underbrace{11\dots1}_p \underbrace{22\dots2}_p \dots \underbrace{99\dots9}_p - \overline{12\dots9}$$

is divisible by p .

Proof. The result is clear for $p = 2$ and for $p = 3$ it follows by computing the sum of digits of the numbers involved, so assume that $p > 3$. By definition we have

$$\underbrace{11\dots1}_p \underbrace{22\dots2}_p \dots \underbrace{99\dots9}_p = 10^{8p} \cdot \frac{10^p - 1}{9} + 2 \cdot 10^{7p} \cdot \frac{10^p - 1}{9} + \dots + 9 \cdot \frac{10^p - 1}{9}.$$

Therefore, we need to prove that

$$10^{8p} \cdot \frac{10^p - 1}{9} - 10^8 + 2 \left(10^{7p} \cdot \frac{10^p - 1}{9} - 10^7 \right) + \dots + 9 \left(\frac{10^p - 1}{9} - 1 \right)$$

is divisible to p . It suffices to check that $10^{kp} \cdot \frac{10^p - 1}{9} \equiv 10^k \pmod{p}$ for all $k \geq 0$. By Fermat's little theorem we have $10^{kp} \equiv 10^k \pmod{p}$, hence it suffices to prove that $10^k \cdot \frac{10^p - 10}{9} \equiv 0 \pmod{p}$. This follows from Fermat's little theorem since $p \neq 3$. \square

2. (Baltic Way 2009) Let p be a prime of the form $6k - 1$ and let a, b, c be integers such that $p \mid a + b + c$ and $p \mid a^4 + b^4 + c^4$. Prove that $p \mid a, b, c$.

Proof. We have $a \equiv -b - c \pmod{p}$ and so

$$b^4 + c^4 + (b + c)^4 \equiv 0 \pmod{p},$$

which can be rewritten as $2(b^2 + bc + c^2)^2 \equiv 0 \pmod{p}$. Thus $p \mid b^2 + bc + c^2$ and then $p \mid b, c$ by corollary 5.30. The result follows. \square

3. (Poland 2010) Let p be an odd prime of the form $3k + 2$. Prove that

$$\prod_{k=1}^{p-1} (k^2 + k + 1) \equiv 3 \pmod{p}.$$

Proof. Since $p \equiv 2 \pmod{3}$, the map $x \mapsto x^3 \pmod{p}$ is a permutation of $\{0, 1, \dots, p-1\}$ (by corollary 5.30) and induces a permutation of $\{2, \dots, p-1\}$. Thus

$$\begin{aligned} (p-2)! \cdot \prod_{k=1}^{p-1} (k^2 + k + 1) &= 3 \cdot \prod_{k=2}^{p-1} (k-1)(k^2 + k + 1) \\ &= 3 \prod_{k=2}^{p-1} (k^3 - 1) \equiv 3 \cdot \prod_{k=2}^{p-1} (k-1) = 3(p-2)! \pmod{p} \end{aligned}$$

and we conclude using the fact that $\gcd((p-2)!, p) = 1$. \square

4. (Iran 2004) Let f be a polynomial with integer coefficients such that for all positive integers m, n there is an integer a such that $n \mid f(a^m)$. Prove that 0 or 1 is a root of f .

Proof. Let p be a prime and choose $n = p$ and $m = p-1$. By hypothesis there is a such that $p \mid f(a^{p-1})$. Since $a^{p-1} \equiv 0, 1 \pmod{p}$, we have $f(a^{p-1}) \equiv f(0), f(1) \pmod{p}$ and so $p \mid f(0)f(1)$. Since this holds for any prime p , we deduce that $f(0)f(1) = 0$, hence the result. \square

5. (Cippola, Rotkiewicz) Prove that if $n_1 > n_2 > \dots > n_k > 1$ are integers with $k > 1$ and $2^{n_k} > n_1$ then $F_{n_1} \dots F_{n_k}$ and $(2^{F_{n_1}} - 1) \dots (2^{F_{n_k}} - 1)$ are pseudo-primes, where $F_n = 2^{2^n} + 1$ is the n th Fermat number.

Proof. Both numbers are clearly composite. Letting $n = F_{n_1} \dots F_{n_k}$, we need to prove that $n \mid 2^{n-1} - 1$. Since F_{n_1}, \dots, F_{n_k} are pairwise relatively prime (see example 3.12), it is enough to prove that $F_{n_i} \mid 2^{n-1} - 1$ for all $1 \leq i \leq k$. Since $F_{n_i} \mid 2^{2^{n_i+1}} - 1$, we are further reduced to $2^{2^{n_i+1}} - 1 \mid$

$2^{n-1} - 1$, or equivalently $2^{n_i+1} \mid n - 1$. Since $F_{n_j} = 2^{2^{n_j}} + 1$ and $2^{n_k} > n_1$ we have $F_{n_j} \equiv 1 \pmod{2^{n_k+1}}$ for all j and so $n - 1 = F_{n_1} \dots F_{n_k} - 1 \equiv 0 \pmod{2^{n_k+1}}$, as desired.

Now let $m = (2^{F_{n_1}} - 1) \dots (2^{F_{n_k}} - 1)$. Again, since F_{n_1}, \dots, F_{n_k} are pairwise relatively prime, so are the numbers $2^{F_{n_1}} - 1, \dots, 2^{F_{n_k}} - 1$ and so it suffices to prove that $2^{F_{n_j}} - 1 \mid 2^{m-1} - 1$ for all j , or equivalently $F_{n_j} \mid m - 1$ for all j . For this, it suffices to prove that $2^{F_{n_u}} - 1 \equiv 1 \pmod{F_{n_j}}$ for all $1 \leq u, j \leq k$, or equivalently $F_{n_j} \mid 2^{F_{n_u}-1} - 1$. Since $F_{n_j} \mid 2^{2^{n_j}+1} - 1$, we are further reduced to $2^{n_j+1} \mid F_{n_u} - 1$ for all j, u , which is again an immediate consequence of the hypothesis that $2^{n_k} > n_1$. \square

6. (India TST 2014) Find all polynomials f with integer coefficients such that $f(n)$ and $f(2^n)$ are relatively prime for all positive integers n .

Proof. If f is constant, then clearly f is a solution of the problem if and only if $f = 1$ or $f = -1$, so assume that f is not constant. Then there is $N > 1$ such that $|f(2^N)| > 1$. Let p be a prime divisor of $f(2^N)$. For any positive integer k we then have $p \mid f(2^N + kp)$ and using the hypothesis of the problem it follows that p does not divide $f(2^{2^N+kp})$. Note that

$$2^{2^N+kp} \equiv 2^{2^N} \cdot 2^k \pmod{p}, \quad \text{hence} \quad f(2^{2^N+kp}) \equiv f(2^{2^N} \cdot 2^k) \pmod{p},$$

thus p does not divide $f(2^{2^N} \cdot 2^k)$. Since $p \mid f(2^N)$ and $2^{2^N} \cdot 2^k - 2^N \mid f(2^{2^N} \cdot 2^k) - f(2^N)$, we deduce that p cannot divide $2^{2^N} \cdot 2^k - 2^N$ for any $k \geq 1$. This is absurd: we can always choose $k \geq 1$ such that $2^N + k \equiv N \pmod{p-1}$ and then $p \mid 2^{2^N} \cdot 2^k - 2^N$ by Fermat's little theorem. Thus no nonconstant polynomial can be solution of the problem and the only solutions are $f = 1$ and $f = -1$. \square

7. (Rotkiewicz) An integer $n > 1$ is called pseudo-prime if n is composite and $n \mid 2^n - 2$. Prove that if p, q are distinct odd primes, then the following statements are equivalent:
- pq is a pseudo-prime.
 - $p \mid 2^{q-1} - 1$ and $q \mid 2^{p-1} - 1$.

c) $(2^p - 1)(2^q - 1)$ is a pseudo-prime.

Proof. Let $n = (2^p - 1)(2^q - 1)$. Suppose that pq is a pseudo-prime and let us prove b). By symmetry, it suffices to prove that $p \mid 2^{q-1} - 1$. But since pq is pseudo-prime, we have $pq \mid 2^{pq} - 2$, thus $2^{pq} \equiv 2 \pmod{p}$. On the other hand, by Fermat's little theorem $2^{pq} \equiv 2^q \pmod{p}$ and so $2^q \equiv 2 \pmod{p}$, as desired. Thus b) holds.

Suppose that b) holds and let us prove that n is a pseudo-prime. Since $2^p - 1$ and $2^q - 1$ are relatively prime, it suffices to prove that each of them divides $2^{n-1} - 1$, or equivalently that p and q divide $n - 1$. But by Fermat's little theorem $n - 1 \equiv 2^q - 1 - 1 = 2^q - 2 \pmod{p}$ and since b) holds we obtain $p \mid n - 1$ and, by symmetry, $q \mid n - 1$, as desired.

Finally, assume that c) holds. Then $2^p - 1 \mid n \mid 2^{n-1} - 1$, thus $p \mid n - 1$. As in the previous paragraph $n - 1 \equiv 2^q - 2 \pmod{p}$ and so $p \mid 2^q - 2$ and, by symmetry, $q \mid 2^p - 2$. Now Fermat's little theorem and what we have already established yield

$$2^{pq} = (2^p)^q \equiv 2^q \equiv 2 \pmod{p}$$

and similarly $2^{pq} \equiv 2 \pmod{q}$, thus $pq \mid 2^{pq} - 2$ and a) is proved. \square

8. (Gazeta Matematica) Find all odd primes p for which $\frac{2^{p-1}-1}{p}$ is a perfect power.

Proof. It is easy to check that $p = 3$ and $p = 7$ are solutions and we will show that these are the only solutions.

Write $2^{p-1} - 1 = px^n$ for some $x, n > 1$. We will discuss two cases.

If n is even, write $n = 2k$ and $z = x^k$, then $(2^{\frac{p-1}{2}} - 1)(2^{\frac{p-1}{2}} + 1) = pz^2$. Since $2^{\frac{p-1}{2}} - 1$ and $2^{\frac{p-1}{2}} + 1$ are relatively prime, we deduce that there is $r \in \{-1, 1\}$ such that

$$2^{\frac{p-1}{2}} + r = u^2, \quad 2^{\frac{p-1}{2}} - r = pv^2$$

for some positive integers u, v with $uv = z$. If $r = 1$, then $u^2 - 1$ is a power of 2, which implies that $u = 3$ and $p = 7$. If $r = -1$, then $2^{\frac{p-1}{2}} \mid u^2 + 1$ and since $u^2 + 1$ cannot be a multiple of 4 we conclude $p \leq 3$, then $p = 3$.

Suppose now that $n > 1$ is odd, then a similar argument yields the existence of $r \in \{-1, 1\}$ and $u, v > 0$ such that $2^{\frac{p-1}{2}} + r = u^n$ and $2^{\frac{p-1}{2}} - r = pv^n$. Then

$$2^{\frac{p-1}{2}} = u^n - r = u^n - r^n = (u - r)(u^{n-1} + \dots + r^{n-1})$$

and $u^{n-1} + \dots + r^{n-1}$ is odd. We deduce that $u^{n-1} + \dots + r^{n-1} = 1$ and this gives $u = 1$, $r = -1$ and $p = 3$. \square

9. (IMO Shortlist 2012) Define $\text{rad}(0) = \text{rad}(1) = 1$ and, for $n \geq 2$ let $\text{rad}(n)$ be the product of the different prime divisors of n . Find all polynomials $f(x)$ with nonnegative integer coefficients such that $\text{rad}(f(n))$ divides $\text{rad}(f(n^{\text{rad}(n)}))$ for all nonnegative integers n .

Proof. Let f be such a polynomial and suppose that f is not the zero polynomial (which is clearly a solution of the problem). Note that $\text{rad}(n^k) = \text{rad}(n)$ for all n and all $k \geq 1$. Let n be a nonnegative integer and define $x_0 = n$ and $x_{k+1} = x_k^{\text{rad}(x_k)}$. Then by assumption $\text{rad}(f(x_k))$ divides $\text{rad}(f(x_{k+1}))$ for all k . On the other hand, $\text{rad}(x_{k+1}) = \text{rad}(x_k) = \text{rad}(n)$ for all k , hence $x_k = n^{\text{rad}(n)^k}$. We conclude that $\text{rad}(f(n))$ divides $\text{rad}(f(n^{\text{rad}(n)^k}))$ for all n and all k .

Since f is not the zero polynomial and since its coefficients are nonnegative, we must have $f(1) \neq 0$. Let p be any prime greater than $f(1)$ and suppose that p divides $f(n)$ for some nonnegative integer n . We will prove that p divides n . Replacing n by $n + pm$ for a suitable m , we may assume that $p - 1$ divides n . Thus for k large enough we have $p - 1 \mid \text{rad}(n)^k$ and since p does not divide n , Fermat's little theorem gives $n^{\text{rad}(n)^k} \equiv 1 \pmod{p}$, thus $f(n^{\text{rad}(n)^k}) \equiv f(1) \pmod{p}$ and so p does not divide $\text{rad}(f(n^{\text{rad}(n)^k}))$. This contradicts the first paragraph and the fact that p divides $f(n)$.

Thus for all primes $p > f(1)$ such that $p \mid f(n)$ for some n we have $p \mid n$. Writing $f(X) = X^k g(X)$ for some nonnegative k and some polynomial g such that $g(0) \neq 0$, we claim that g is constant. Indeed, otherwise by Schur's theorem 4.67 there are infinitely many primes p for which $p \mid g(n)$ for some n . By the previous paragraph each such p divides n and so it also divides $g(0)$. But then $g(0) = 0$, a contradiction. Hence g is constant and $f(X) = cX^k$ for some nonnegative integer c . Since all these polynomials are clearly solutions of the problem, we are done. \square

10. (Turkey TST 2013) Find all pairs of positive integers (m, n) such that

$$2^n + (n - \varphi(n) - 1)! = n^m + 1.$$

Proof. Let n, m be a solution of the problem, then clearly $n > 1$. Let p be the smallest prime divisor of n , then

$$n - \varphi(n) \geq \frac{n}{p}$$

since all multiples of p between 1 and n are not relatively prime to n .

If $p \leq n - \varphi(n) - 1$ then taking the original equation modulo p yields $p \mid 2^n - 1$. Since $p \mid 2^{p-1} - 1$ we obtain $p \mid 2^{\gcd(n, p-1)} - 1 = 1$, the last equality being a consequence of the fact that p is the smallest prime factor of n , thus $\gcd(n, p-1) = 1$. Hence we must have $p \geq n - \varphi(n) \geq \frac{n}{p}$, that is $n \leq p^2$. Since p is the smallest prime factor of n , we deduce that $n = p$ or $n = p^2$. If $n = p$ the equality becomes $2^n = n^m$ which forces $n = 2$ and then $m = 2$. Suppose that $n = p^2$ for a prime p , then the equation becomes $2^{p^2} + (p-1)! = p^{2m} + 1$. If $p > 3$ then the left-hand side is a multiple of 4, while the right-hand side is not. Thus $p \leq 3$. For $p = 2$ we obtain $m = 2$ and for $p = 3$ we obtain $513 = 9^m$, with no solution. We conclude that the solutions of the problem are $(m, n) = (2, 2), (2, 4)$. \square

11. (Serbia 2015) Find all nonnegative integers x, y such that

$$(2^{2015} + 1)^x + 2^{2015} = 2^y + 1.$$

Proof. There are two obvious solutions, namely $(x, y) = (0, 2015)$ and $(x, y) = (1, 2016)$. Assume now that (x, y) is a solution with $x > 1$. Then $2^y > (2^{2015} + 1)^2$, so $y > 4030$. Taking the equation mod 2^{4030} yields

$$1 + 2^{2015}x + 2^{2015} \equiv 1 + 2^y \equiv 1 \pmod{2^{4030}},$$

which shows in particular that $16 \mid x + 1$. Next, we work modulo 17. We have $2^4 \equiv -1 \pmod{17}$, thus

$$2^{2015} \equiv 2^{4 \cdot 503 + 3} \equiv -8 \equiv 9 \pmod{17},$$

thus we obtain $10^x + 8 \equiv 2^y \pmod{17}$. Since $x \equiv -1 \pmod{16}$, we have by Fermat's little theorem

$$10^x \equiv 10^{-1} \equiv -5 \pmod{17}.$$

We conclude that $2^y \equiv 3 \pmod{17}$. Writing $y = 4k + r$ with $0 \leq r \leq 3$, we finally obtain $(-1)^k \cdot 2^r \equiv 3 \pmod{17}$. A simple verification shows that there are no such r, k , showing therefore that there are no solutions with $x > 1$. Therefore we have already found all solutions. \square

12. (Italy 2010) If n is a positive integer, let

$$a_n = 2^{n^3+1} - 3^{n^2+1} + 5^{n+1}.$$

Prove that infinitely many primes divide at least one of the numbers a_1, a_2, \dots

Proof. Suppose that this is not the case and let p_1, \dots, p_k be all odd primes dividing at least one of the numbers a_1, a_2, \dots . Let

$$n = s(p_1 - 1) \dots (p_k - 1)$$

for some positive integer s . Note that since 5 is among p_1, \dots, p_k (as $5 \mid a_1$) we have $4 \mid n$ and so $a_n \equiv 2 \pmod{4}$, $a_n \equiv 1 \pmod{3}$ and $a_n \equiv 2^{n^3+1} + 2^{n^2+1} \equiv 4 \pmod{5}$. In particular if $a_n > 2$ (which is definitely the case for s large enough, actually even for any $s \geq 1$) then

a_n must have a prime factor p greater than 5. By assumption this prime factor is among p_1, \dots, p_k and so $p - 1 \mid n$. Using Fermat's little theorem we obtain $a_n \equiv 2 - 3 + 5 \equiv 4 \pmod{p}$, a contradiction. \square

13. (China TST 2010) Find all positive integers $m, n \geq 2$, such that

- a) $m + 1$ is a prime number of the form $4k - 1$;
- b) there is a prime number p and a nonnegative integer a such that

$$\frac{m^{2^n-1} - 1}{m - 1} = m^n + p^a.$$

Proof. Let $q = m + 1$ and note that by assumption $m \equiv 2 \pmod{4}$. Taking the equation mod 4 yields $3 \equiv q \equiv p^a \pmod{4}$, which forces a being odd and $p \equiv 3 \pmod{4}$.

Next, write the equation as

$$\frac{m^{2^n} - 1}{m - 1} - (m^{n+1} + 1) = mp^a.$$

Assume that $n + 1$ is even, and let $v_2(n + 1) = r$, then clearly $r < n$ so that $m^{2^r} + 1$ divides both $\frac{m^{2^n}-1}{m-1}$ and $m^{n+1} + 1$. Thus $m^{2^r} + 1 \mid mp^a$ and since $\gcd(m^{2^r} + 1, m) = 1$ we deduce that $m^{2^r} + 1$ is a power of p . But then $p \mid m^{2^r} + 1$, a contradiction with $p \equiv 3 \pmod{4}$.

Hence $n + 1$ is odd. Then $q = m + 1$ divides both $\frac{m^{2^n}-1}{m-1}$ and $m^{n+1} + 1$, hence as before $q \mid p^a$. This forces $q = p$. Now, write the equation as

$$q + m^2 + m^3 + \dots + m^{2^n-2} = m^n + q^a.$$

If $n \geq 3$ then the left-hand side is congruent to $q + 4 \pmod{8}$, while the right-hand side is congruent to $q^a \equiv q \pmod{8}$, a contradiction. Hence we must have $n = 2$ and the equation reduces to $m + 1 = p^a$. We must have $a = 1$, hence the answer is given by pairs $(m, n) = (q - 1, 2)$ with q a prime of the form $4k - 1$. \square

14. Let p be a prime. Prove that there is a positive integer n such that p is the smallest prime divisor of $n! + 1$.

Proof. Simply choose $n = p - 1$. By Wilson's theorem $p \mid n! + 1$, and if q is a prime divisor of $n! + 1$ then clearly $q > n = p - 1$ (otherwise $q \mid n!$, a contradiction with $q \mid n! + 1$) and so $q \geq p$. \square

15. Let $n > 1$ and suppose that there is $k \in \{0, 1, \dots, n - 1\}$ such that

$$k!(n - k - 1)! + (-1)^k \equiv 0 \pmod{n}.$$

Prove that n is a prime.

Proof. Suppose that n is composite and let p be its smallest prime factor. Then $p \leq \sqrt{n}$. On the other hand by assumption p does not divide $k!(n - k - 1)!$, thus $p \geq k + 1$ and $p \geq n - k$. We deduce that $2p \geq n + 1 \geq p^2 + 1$, that is $(p - 1)^2 \leq 0$, a contradiction. Hence n is a prime number. \square

16. For each positive integer n find the greatest common divisor of $n! + 1$ and $(n + 1)!$.

Proof. Suppose that a prime p divides $n! + 1$ and $(n + 1)!$. Then $p \leq n + 1$ (since it divides $(n + 1)!$) and $p > n$ (otherwise p divides $n!$, hence cannot divide $n! + 1$). Thus $p = n + 1$. This shows that if $n + 1$ is not a prime, then $\gcd(n! + 1, (n + 1)!) = 1$, while if $n + 1 = p$ is a prime, then $\gcd(n! + 1, (n + 1)!) is a power of p . By Wilson's theorem, p divides $n! + 1 = (p - 1)! + 1$. Hence p divides $\gcd(n! + 1, (n + 1)!)$. Moreover, p^2 does not divide $(n + 1)!$, since otherwise p would divide $n!$ and so $n + 1 = p \leq n$, a contradiction. Therefore $\gcd(n! + 1, (n + 1)!) = p$ if $n + 1 = p$ is a prime. $\square$$

17. Let p be a prime and let a_1, a_2, \dots, a_{p-1} be consecutive integers.

- a) What are the possible remainders of $a_1 a_2 \dots a_{p-1}$ when divided by p ?
b) Suppose that $p \equiv 3 \pmod{4}$. Prove that a_1, \dots, a_{p-1} cannot be partitioned into two sets with the same product of their elements.

Proof. a) If one of the a_i 's is divisible by p , then clearly the remainder of $a_1 a_2 \dots a_{p-1}$ when divided by p is 0, so assume that a_1, \dots, a_{p-1} are not divisible by p . Since a_1, \dots, a_{p-1} are $p-1$ consecutive integers, they give pairwise different remainders when divided by p , and by assumption these remainders are nonzero. It follows that the remainders of a_1, \dots, a_{p-1} when divided by p are a permutation of $1, 2, \dots, p-1$ and so by Wilson's theorem

$$a_1 a_2 \dots a_{p-1} \equiv 1 \cdot 2 \cdot \dots \cdot (p-1) = (p-1)! \equiv -1 \pmod{p}.$$

Hence the answer of the problem is given by 0 and $p-1$.

b) Suppose that there is such a partition in two sets A and B and let x, y be the product of the elements of A , respectively B . By assumption $x = y$, thus

$$x^2 = xy = a_1 a_2 \dots a_{p-1}.$$

If $a_1 a_2 \dots a_{p-1}$ is not a multiple of p , then part a) yields $x^2 \equiv -1 \pmod{p}$, thus $p \mid x^2 + 1$, contradicting the fact that $p \equiv 3 \pmod{4}$. Thus $a_1 \dots a_{p-1}$ is a multiple of p and so one of the a_i 's is a multiple of p . Since a_1, \dots, a_{p-1} are $p-1$ consecutive integers, exactly one of them is a multiple of p . But then exactly one of the numbers x and y is a multiple of p , contradicting the equality $x = y$. \square

18. Find two primes p such that $(p-1)! + 1 \equiv 0 \pmod{p^2}$.

Proof. Clearly $p = 5$ is a solution of the problem. One can tediously check that $p = 13$ is also a solution. Here is a rather tricky way to do it without actually computing $12!$: we have

$$12! = (2 \cdot 3 \cdot 4 \cdot 7) \cdot (6 \cdot 11 \cdot 8 \cdot 2 \cdot 4) \cdot (5 \cdot 9 \cdot 5 \cdot 3)$$

and one easily checks that the products in each parenthesis are congruent to $-1 \pmod{169}$: for the second set of parentheses note that

$$6 \cdot 11 \cdot 8 \cdot 2 \cdot 4 = 64 \cdot 66 = 65^2 - 1 \equiv -1 \pmod{13^2}$$

and similarly for the third one note that

$$5 \cdot 9 \cdot 5 \cdot 3 = 25 \cdot 27 = 26^2 - 1 \equiv -1 \pmod{13^2}.$$

The only known such primes are 5, 13, 563, all others are $> 2 \cdot 10^{13}$. \square

19. Find all sequences a_1, a_2, \dots of positive integers such that for all positive integers m, n

$$m! + n! \mid a_m! + a_n!.$$

Proof. Clearly $n! \mid a_n!$ for all n , thus $a_n \geq n$ for all $n \geq 1$. Next, if p is a prime then Wilson's theorem gives $p \mid 1! + (p-1)! \mid a_1! + a_{p-1}!$. If $p > a_1$, then p cannot divide $a_1!$ and so p cannot divide $a_{p-1}!$. Since $a_{p-1} \geq p-1$, we must have $a_{p-1} = p-1$ for all $p > a_1$. Thus for all $n \geq 1$ and all primes $p > a_1$ we have $(p-1)! + n! \mid (p-1)! + a_n!$, thus also $(p-1)! + n! \mid a_n! - n!$. Fixing $n \geq 1$ and varying p yields $a_n! = n!$. We conclude that the only solution of the problem is given by the sequence $a_n = n$ for all $n \geq 1$. \square

20. Let p be an odd prime. A subset A of \mathbf{Z} is called a complete set of nonzero residue classes modulo p if A consists of $p-1$ integers giving pairwise distinct and nonzero remainders when divided by p . Prove that if $A = \{a_1, a_2, \dots, a_{p-1}\}$ and $B = \{b_1, b_2, \dots, b_{p-1}\}$ are complete sets of nonzero residue classes modulo p , then $\{a_1 b_1, \dots, a_{p-1} b_{p-1}\}$ is not a complete set of nonzero residue classes.

Proof. Using problem 17 it follows that for any complete set A of nonzero residue classes modulo p we have $\prod_{a \in A} a \equiv -1 \pmod{p}$. Therefore

$$\prod_{i=1}^{p-1} a_i \equiv \prod_{i=1}^{p-1} b_i \equiv -1 \pmod{p}$$

and consequently

$$\prod_{i=1}^{p-1} (a_i b_i) \equiv 1 \not\equiv -1 \pmod{p}.$$

It follows that $\{a_1b_1, \dots, a_{p-1}b_{p-1}\}$ is not a complete set of nonzero residue classes. \square

21. (Clement's criterion) Let n be an integer greater than 2. Prove that n and $n + 2$ are both primes if and only if

$$4((n-1)! + 1) + n \equiv 0 \pmod{n(n+2)}.$$

Proof. Suppose that n and $n + 2$ are both primes. By Wilson's theorem $4((n-1)! + 1) + n \equiv 0 \pmod{n}$. By theorem 5.49 we have $2(n-1)! \equiv -1 \pmod{n+2}$, thus

$$4((n-1)! + 1) + n \equiv -2 + 4 + n \equiv 0 \pmod{n+2}.$$

This proves one implication, since $\gcd(n, n+2) = 1$.

Assume now that $4((n-1)! + 1) + n \equiv 0 \pmod{n(n+2)}$, in particular $n \mid 4((n-1)! + 1)$. Let us prove that n must be odd. Write $v_2(x)$ for the exponent of 2 in the prime factorization of x . If n is even, then $v_2(n) \leq v_2(4((n-1)! + 1)) = 2$. Moreover, if d is the largest odd divisor of n , then $d < n$ so $d \mid (n-1)!$ and since $d \mid n \mid 4((n-1)! + 1)$ we deduce that $d \mid 4$ and then $d = 1$, that is n is a power of 2. Since $v_2(n) \leq 2$, it follows that $n = 2$ or $n = 4$, but neither of these satisfies $4((n-1)! + 1) + n \equiv 0 \pmod{n(n+2)}$.

Thus n is odd, and since $n \mid 4((n-1)! + 1)$ we deduce that $n \mid (n-1)! + 1$ and then that n is a prime. Next, let $k = n + 2$, then

$$k \mid 4((k-3)! + 1) + k - 2 = 4(k-3)! + k + 2,$$

thus $k \mid 4(k-3)! + 2$. Since k is odd, we deduce that $k \mid 2(k-3)! + 1$. But then

$$(k-1)! = (k-3)!(k-2)(k-1) \equiv 2(k-3)! \equiv -1 \pmod{k},$$

thus k is a prime and we are done. \square

22. Let $n > 1$ be an integer. Prove that there exists a positive integer k and $\varepsilon \in \{-1, 1\}$ such that $2k + 1 \mid n + \varepsilon k!$.

Proof. Let p be an odd prime divisor of $n^2 + 1$ (such p exists since $n^2 + 1 > 4$ and is not divisible by 4). Then $p \equiv 1 \pmod{4}$ so taking $k = \frac{p-1}{2}$ we have $p = 2k + 1 \mid (k!)^2 + 1$ and $p \mid n^2 + 1$, hence $p \mid (k!)^2 - n^2$. Since p is a prime, $p = 2k + 1$ divides one of the numbers $n + k!$ and $n - k!$, and the result follows. \square

23. (Moldova TST 2007) Prove that for infinitely many prime numbers p there is a positive integer n such that n does not divide $p-1$ and $p \mid n!+1$.

Proof. The key is the congruence

$$k!(p-1-k)! \equiv (-1)^{k-1} \pmod{p}$$

established in theorem 5.49. Take for now an arbitrary even number k and assume that $p \mid k! - 1$, then necessarily $p > k$ (otherwise $p \mid k!$) and the above congruence shows that $p \mid (p-k-1)! + 1$. We want to ensure that $p-k-1$ does not divide $p-1$. If $p-k-1 \mid p-1$, then $p-1-k \mid k$. To make our life easier, choose $k = 2q$ with $q > 2$ a prime, then necessarily $p > 3$ (since $p > k$) and so $p-1-k$ (which is assumed to divide $k = 2q$) equals 2 or $2q$. We cannot have $p-1-k = 2$ since in this case $p \mid 2! + 1 = 3$, thus $p-1-2q = 2q$ and so $p = 1 + 4q$. It is now very easy to conclude: for each odd prime q choose a prime divisor p of $(2q)! - 1$ of the form $4k + 3$ (this is possible since $(2q)! - 1 \equiv 3 \pmod{4}$), then by the above discussion p is a solution of the problem. Since $p > 2q$, when q varies over all odd primes we get infinitely many solutions of the problem. \square

24. Find all polynomials f with integer coefficients such that for all primes p we have $f(p) \mid (p-1)! + 1$.

Proof. If f is constant, say $f = c$, then $c \mid 2$ and $c \mid 3$ thus $c = \pm 1$. Conversely, the constant polynomials ± 1 are solutions of the problem,

so assume now that f is not constant. Replacing f with $-f$ we may assume that the leading coefficient of f is positive. Thus if p is a large enough prime, then $f(p) > 1$. Let q be a prime factor of $f(p)$ and assume that $q \neq p$. By Dirichlet's theorem there are infinitely many primes $r \equiv p \pmod{q}$. Then $f(r) \equiv f(p) \equiv 0 \pmod{q}$ and so $q \mid f(r) \mid (r-1)! + 1$, yielding $q \geq r$ for infinitely many primes r , a contradiction. Hence if p is a large enough prime, then $f(p)$ is a power of p , say $f(p) = p^{a_p}$ for some positive integer a_p . If $d = \deg f$, then there is a constant $c > 0$ such that $f(x) < cx^d$ for all $x > 1$. Thus $p^{a_p} < cp^d$ for all large enough p and so the sequence (a_p) is bounded. It follows that there is a positive integer a such that $f(p) = p^a$ for infinitely many primes p , but then $f(x) = x^a$ for all x . We deduce that $p^a \mid (p-1)! + 1$ for all primes p and so $2^a \mid 2$. But then $a = 1$ and $f(x) = x$. Hence the solutions of the problem are the polynomials $\pm 1, \pm X$. \square

25. (adapted from Serbia 2010) Let a, n be positive integers such that $a > 1$ and $a^n + a^{n-1} + \dots + a + 1$ divides $a^{n!} + a^{(n-1)!} + \dots + a^{1!} + 1$. Prove that $n = 1$ or $n = 2$.

Proof. In general, suppose that $a^n + a^{n-1} + \dots + a + 1$ divides $a^{k_1} + \dots + a^{k_n} + 1$ for some positive integers k_i . Let r_i be the remainder of k_i modulo $n + 1$. Since $a^n + \dots + a + 1 \mid a^{n+1} - 1$, it follows that $a^n + \dots + a + 1 \mid a^{r_1} + \dots + a^{r_n} + 1$. Suppose that c_i is the number of r_j 's equal to j . Thus $c_0 + \dots + c_n = n$ and $a^n + \dots + a + 1 \mid c_n a^n + \dots + c_1 a + c_0 + 1$.

Consider now an $n + 1$ -tuple (b_0, \dots, b_n) of nonnegative integers which minimizes $b_0 + \dots + b_n$ subject to the condition $a^n + \dots + a + 1 \mid b_n a^n + \dots + b_1 a + b_0$. If some $b_i \geq a$, then replacing b_i by $b_i - a$ and b_{i+1} by $b_{i+1} + 1$ does not change $b_n a^n + \dots + b_1 a + b_0$, but diminishes $b_0 + \dots + b_n$ and yields an $n + 1$ -tuple with a smaller sum, a contradiction. Hence all b_i are less than a and so $b_n a^n + \dots + b_1 a + b_0 = r(a^n + \dots + a + 1)$ for some $r < a$. But uniqueness of base a representation then yields $b_n = \dots = b_0 = r$. We conclude that the minimal value of $b_0 + \dots + b_n = n + 1$ and is obtained only when $b_0 = \dots = b_n = 1$.

Combining the previous paragraphs, we deduce that $c_n = \dots = c_0 = 1$ and so r_1, \dots, r_n are a permutation of $1, 2, \dots, n$. Applying this to our initial problem, we deduce that $1!, 2!, \dots, n!$ give remainders $1, 2, \dots, n$ (in some order) when divided by $n + 1$. If $n + 1$ is a composite number and $n > 3$, then $n!$ gives remainder 0 when divided by $n + 1$, hence this case is impossible. If $n + 1 = p$ is a prime and $p > 3$, then Wilson's theorem gives $(p-2)! \equiv 1 \pmod{p}$, hence $(n-1)!$ and $1!$ give the same remainder mod $n + 1$, again a contradiction. Hence necessarily $n = 1$ or $n = 2$. \square

26. Let p be a prime. Prove that the sequence of remainders mod p of the numbers $1, 2^2, 3^3, 4^4, \dots$ is periodic and find its least period.

Proof. To prove periodicity, note that

$$(n + p^2 - p)^{n+p^2-p} \equiv n^{n+p^2-p} \equiv n^n \pmod{p},$$

the last congruence being a consequence of Fermat's little theorem. We need to find the smallest positive integer T for which $(n + T)^{n+T} \equiv n^n \pmod{p}$ for all n . Taking $n = p$ we obtain $(p + T)^{p+T} \equiv p^p \pmod{p}$, which reduces to $T \equiv 0 \pmod{p}$. But then $n^n \equiv (n + T)^{n+T} \equiv n^{n+T} \pmod{p}$ for all n . Hence $p \mid n^n(n^T - 1) \pmod{p}$. It follows that $n^T \equiv 1 \pmod{p}$ for all n relatively prime to p and so (by corollary 5.76) $p-1 \mid T$. Conversely, if $p-1 \mid T$, then $p \mid n^n(n^T - 1)$ for all n , by Fermat's little theorem. We conclude that T is a period of the sequence if and only if $p(p-1) \mid T$. Thus the smallest period is $p(p-1)$. \square

27. (Don Zagier) Somebody incorrectly remembered Fermat's little theorem as saying that the congruence $a^{n+1} \equiv a \pmod{n}$ holds for all integers a . Describe the set of integers n for which this property is in fact true.

Proof. The answer is 1, 2, 6, 42, 1806. Let n be such an integer, then n is squarefree: if $p \mid n$ we can choose $a = p$ and obtain $n \mid p(p^n - 1)$, thus p^2 does not divide n . So, assuming that $n > 1$ (note that $n = 1$ is clearly a solution), we can write $n = p_1 \dots p_k$ with $p_1 < \dots < p_k$ distinct primes. Fix $i \in \{1, 2, \dots, k\}$. By assumption for any a which is not a multiple of

p_i we have $p_i \mid n \mid a(a^n - 1)$, thus $p_i \mid a^n - 1$. We deduce from corollary 5.76 that $p_i - 1 \mid n$. In particular $p_1 - 1 \mid n$ and since p_1 is the smallest prime factor of n , this forces $p_1 - 1 = 1$ and so $p_1 = 2$. Similarly, from $p_2 - 1 \mid n$ and $\gcd(p_2 - 1, p_2 \dots p_k) = 1$ we obtain $p_2 - 1 \mid 2$ and then $p_2 = 3$ (if $k \geq 2$). Continuing like this we obtain $p_3 = 7$ (if $k \geq 3$), then $p_4 = 43$ (if $k \geq 4$). Assuming that $k \geq 5$ we obtain $p_5 - 1 \mid 1806$. It is however easy to see that there is no such prime p_5 , thus $k \leq 4$ and the solutions are given by $1, 2, 2 \cdot 3, 2 \cdot 3 \cdot 7, 2 \cdot 3 \cdot 7 \cdot 43$, i.e. $1, 2, 6, 42, 1806$. It is not difficult to see, using Fermat's little theorem, that if n is squarefree and $p - 1 \mid n$ for all primes $p \mid n$, then n is a solution of the problem. \square

28. Let p be an odd prime. Find the largest degree of a polynomial f with the following properties:

- a) $\deg f < p$.
- b) the coefficients of f are integers between 0 and $p - 1$.
- c) If m, n are integers and p does not divide $m - n$, then p does not divide $f(m) - f(n)$.

Proof. First of all, the polynomial $f(X) = X^{p-2}$ is a solution of the problem. Indeed, assuming that $p \mid m^{p-2} - n^{p-2}$ and that p does not divide $m - n$, we see that p does not divide mn . But then using Fermat's little theorem we obtain $m^{-1} \equiv n^{-1} \pmod{p}$ (where x^{-1} is the inverse of x modulo p) and then $m \equiv n \pmod{p}$, a contradiction.

We will now prove that there is no polynomial f of degree $p - 1$ satisfying a), b) and c). Suppose that such a polynomial exists, and write

$$f = cX^{p-1} + g$$

with $c \in \{1, 2, \dots, p - 1\}$ and $\deg(g) \leq p - 2$. By condition c)

$$f(0) + f(1) + \dots + f(p-1) \equiv 0 + 1 + \dots + (p-1) = \frac{p(p-1)}{2} \equiv 0 \pmod{p}.$$

On the other hand Fermat's little theorem combined with corollary 5.77 yield

$$\begin{aligned} f(0) + f(1) + \dots + f(p-1) &= c(1^{p-1} + 2^{p-1} + \dots + (p-1)^{p-1}) \\ &\quad + g(0) + g(1) + \dots + g(p-1) \\ &\equiv c(1 + 1 + \dots + 1) + 0 \equiv -c \pmod{p}, \end{aligned}$$

a contradiction since c is not a multiple of p . Thus the answer of the problem is $p-2$. \square

29. (Iran TST 2012) Let $p > 2$ be an odd prime. If $i \in \{0, 1, \dots, p-1\}$ and $f = a_0 + a_1X + \dots + a_nX^n$ is a polynomial with integer coefficients, we say that f is i -remainder if

$$\sum_{j>0, p-1|j} a_j \equiv i \pmod{p}.$$

Prove that the following statements are equivalent:

- a) f, f^2, \dots, f^{p-2} are 0-remainder and f^{p-1} is 1-remainder.
- b) $f(0), f(1), \dots, f(p-1)$ form a complete residue system modulo p .

Proof. Corollary 5.77 shows that for any polynomial

$$f = a_0 + a_1X + \dots + a_nX^n$$

with integer coefficients we have (with the convention $0^0 = 1$)

$$\begin{aligned} f(0) + f(1) + \dots + f(p-1) &= \sum_{j=0}^n a_j(0^j + 1^j + \dots + (p-1)^j) \\ &\equiv - \sum_{j>0, p-1|j} a_j \pmod{p}, \end{aligned}$$

hence f is i -remainder if and only if

$$f(0) + f(1) + \dots + f(p-1) \equiv -i \pmod{p}.$$

Suppose first that $f(0), f(1), \dots, f(p-1)$ form a complete residue system modulo p . We need to prove that $f(0)^j + \dots + f(p-1)^j \equiv 0 \pmod{p}$ for $1 \leq j \leq p-2$ and $f(0)^{p-1} + \dots + f(p-1)^{p-1} \equiv -1 \pmod{p}$. But since $f(0)^j + f(1)^j + \dots + f(p-1)^j \equiv 0^j + 1^j + \dots + (p-1)^j \pmod{p}$, the result follows from corollary 5.77.

Conversely, assume that f, f^2, \dots, f^{p-2} are 0-remainder and f^{p-1} is 1-remainder and let $a_j = f(j)$. By assumption $a_0^j + \dots + a_{p-1}^j \equiv 0 \pmod{p}$ for $1 \leq j \leq p-2$ and $a_0^{p-1} + \dots + a_{p-1}^{p-1} \equiv -1 \pmod{p}$. This last congruence combined with Fermat's little theorem shows that exactly one of a_0, \dots, a_{p-1} is a multiple of p . Similarly, if c is any integer then

$$\begin{aligned} \sum_{k=0}^{p-1} (a_k - c)^{p-1} &= \sum_{k=0}^{p-1} a_k^{p-1} - \binom{p-1}{1} c \sum_{k=0}^{p-1} a_k^{p-2} + \dots + p(-c)^{p-1} \\ &\equiv -1 \pmod{p} \end{aligned}$$

and so exactly one of $a_0 - c, \dots, a_{p-1} - c$ is a multiple of p . This is equivalent to the fact that a_0, \dots, a_{p-1} is a complete residue system modulo p . \square

30. Find all integers $n > 2$ for which $n \mid 2^n + 3^n + \dots + (n-1)^n$.

Proof. Let n be such a number and let p be a prime divisor of n . Write $n = kp$ and observe that

$$1^n + 2^n + \dots + (n-1)^n \equiv k(1^n + 2^n + \dots + (p-1)^n) \pmod{p},$$

since

$$(px+1)^n + \dots + (px+p-1)^n \equiv 1^n + \dots + (p-1)^n \pmod{p}$$

for all x . We deduce that

$$k(1^n + 2^n + \dots + (p-1)^n) \equiv 1 \pmod{p}.$$

In particular p does not divide k , i.e. p^2 does not divide n . Since p was an arbitrary prime, this means that n is squarefree. On the other

hand, the sum $1^n + 2^n + \dots + (p-1)^n$ is not divisible by p thanks to the previous congruence. Corollary 5.77 implies therefore that $p-1 \mid n$ and so $1^n + \dots + (p-1)^n \equiv -1 \pmod{p}$. Thus $k \equiv -1 \pmod{p}$. In other words, n must be squarefree, $p-1 \mid n$ and $p \mid \frac{n}{p} + 1$ for all $p \mid n$. The solution of problem 27 shows that the first conditions are already enough to ensure that $n = 6, 42$ or 1806 . It is not difficult to check that each of these numbers is a solution of the problem: by the above computations this comes down to checking that $p-1 \mid n$ and $p \mid \frac{n}{p} + 1$ for each $p \mid n$, which is straightforward. Thus the solutions of the problem are 6, 42 and 1806. \square

31. (Alon, Dubiner) Let p be a prime and let $a_1, \dots, a_{3p}, b_1, \dots, b_{3p}$ be integers such that

$$\sum_{i=1}^{3p} a_i \equiv \sum_{i=1}^{3p} b_i \equiv 0 \pmod{p}.$$

Prove that there is a subset $I \subset \{1, 2, \dots, 3p\}$ with p elements such that

$$\sum_{i \in I} a_i \equiv \sum_{i \in I} b_i \equiv 0 \pmod{p}.$$

Proof. Consider the following system of congruences

$$\begin{aligned} \sum_{i=1}^{3p-1} a_i x_i^{p-1} &\equiv 0 \pmod{p}, \\ \sum_{i=1}^{3p-1} b_i x_i^{p-1} &\equiv 0 \pmod{p}, \\ \sum_{i=1}^{3p-1} x_i^{p-1} &\equiv 0 \pmod{p}. \end{aligned}$$

Since $3(p-1) < 3p-1$ and $x_1 = x_2 = \dots = x_{3p-1} = 0$ is a solution, by corollary 5.88 the system has a nontrivial solution $(x_i)_{1 \leq i \leq 3p-1}$. Let

$$I = \{i \mid 1 \leq i \leq 3p-1, x_i \neq 0 \pmod{p}\},$$

then Fermat's little theorem combined with the equations of the system yields

$$\sum_{i \in I} a_i \equiv 0 \pmod{p}, \quad \sum_{i \in I} b_i \equiv 0 \pmod{p}.$$

Moreover, we have $p \mid |I|$, thus $|I| = p$ or $|I| = 2p$. In the first case it suffices to choose the set I , in the second case we can choose its complement (this is where the hypothesis that the sum of the a_i 's and the sum of the b_j 's are multiples of p is used). \square

32. Prove that for any $n > 1$ the number $\binom{n}{0}^4 + \binom{n}{1}^4 + \dots + \binom{n}{n}^4$ is a multiple of any prime $p \in (n, \frac{4}{3}n]$.

Proof. Let

$$A = (n+1)(n+2)\dots(p-1),$$

with the convention that $A = 1$ if $p = n+1$. For all $j \in \{0, 1, \dots, n\}$ we have

$$\begin{aligned} A \binom{n}{j} &= \frac{(p-1) \cdot \dots \cdot n(n-1)\dots(n-j+1)}{j!} \\ &= (n-j+1) \cdot \dots \cdot (p-1-j) \cdot \frac{(p-j)(p-j+1)\dots(p-1)}{j!}. \end{aligned}$$

Since $(p-j)(p-j+1)\dots(p-1)$ is congruent to $(-1)^j \cdot j!$ modulo p , we deduce that

$$A^4 \cdot \sum_{j=0}^n \binom{n}{j}^4 \equiv \sum_{j=0}^n f(j) \pmod{p},$$

where

$$f(X) = (n-X+1)^4 \cdot \dots \cdot (p-X-1)^4.$$

Note that

$$\deg(f) = 4(p-1-n) < p-1$$

thanks to the hypothesis of the problem. Also

$$f(n+1) = \dots = f(p-1) = 0,$$

thus (using corollary 5.77)

$$\sum_{j=0}^n f(j) \equiv \sum_{j=0}^{p-1} f(j) \equiv 0 \pmod{p}$$

and we conclude that $A^4 \cdot \sum_{j=0}^n \binom{n}{j}^4$ is a multiple of p . Since A is not a multiple of p , the result follows. \square

33. Let f be a monic polynomial of degree $n \geq 1$, with integer coefficients. Suppose that b_1, \dots, b_n are pairwise distinct integers and that for infinitely many primes p the simultaneous congruences

$$f(x + b_1) \equiv f(x + b_2) \equiv \dots \equiv f(x + b_n) \equiv 0 \pmod{p}$$

have a common solution. Prove that the equations

$$f(x + b_1) = \dots = f(x + b_n) = 0$$

have a common integral solution.

Proof. Since b_1, \dots, b_n are pairwise distinct, they are pairwise distinct modulo p for all sufficiently large primes p (more precisely for all primes p not dividing the nonzero integer $\prod_{i \neq j} (b_i - b_j)$). We only consider such primes p from now on. By assumption for infinitely many such primes p we can find an integer x_p such that $f(x_p + b_i) \equiv 0 \pmod{p}$ for all $1 \leq i \leq n$. Using Lagrange's theorem, we deduce that

$$f(X) \equiv \prod_{i=1}^n (X - x_p - b_i) \pmod{p},$$

since the difference between the two sides is a polynomial of degree at most $n - 1$ that has at least n distinct roots modulo p (namely the numbers $x_p + b_i$ with $1 \leq i \leq n$). Writing

$$f(X) = X^n + a_{n-1}X^{n-1} + \dots + a_0,$$

we deduce that

$$a_{n-1} \equiv -nx_p - \sum_{i=1}^n b_i \pmod{p}.$$

Letting

$$A = -a_{n-1} - \sum_{i=1}^n b_i,$$

we obtain $nx_p \equiv A \pmod{p}$. Since $nx_p \equiv A \pmod{p}$ and $f(x_p + b_i) \equiv 0 \pmod{p}$, it follows that for all $1 \leq i \leq n$

$$n^d f\left(\frac{A}{n} + b_i\right) \equiv 0 \pmod{p}.$$

The left-hand side is independent of p and since the congruence holds for infinitely many primes we deduce that $f\left(\frac{A}{n} + b_i\right) = 0$ for $1 \leq i \leq n$.

By the rational root theorem the rational number $\frac{A}{n} + b_i$ must be an integer, thus $x := \frac{A}{n}$ is an integer and $f(x + b_i) = 0$ for $1 \leq i \leq n$. The result follows. \square

34. (Romania TST 2016) Given a prime p , prove that

$$\sum_{k=1}^{\left\lfloor \frac{q}{p} \right\rfloor} k^{p-1}$$

is not divisible by q for all but finitely many primes q .

Proof. Since each prime $q \neq p$ is of the form $q = pn + r$ for some $1 \leq r \leq p-1$, it will be enough to prove that for each such r there are only finitely many n such that $pn + r \mid 1^{p-1} + 2^{p-1} + \dots + n^{p-1}$. Let us fix $r \in \{1, 2, \dots, p-1\}$ and assume that $pn + r \mid 1^{p-1} + \dots + n^{p-1}$ for infinitely many n .

The first key observation is that for each $k \geq 1$ there is a polynomial f_k with rational coefficients, of degree $k+1$ and leading coefficient $\frac{1}{k+1}$, such that

$$1^k + 2^k + \dots + n^k = f_k(n)$$

for all $n \geq 1$. This is easily established by induction on k , using the following relation to pass from $k-1$ to k

$$(n+1)^{k+1} - 1 = \binom{k+1}{1}(1^k + \dots + n^k) + \binom{k+1}{2}(1^{k-1} + \dots + n^{k-1}) + \dots + n.$$

This relation follows immediately by adding up the relations (for $1 \leq j \leq n$)

$$(j+1)^{k+1} - j^{k+1} = \binom{k+1}{1}j^k + \binom{k+2}{2}j^{k-1} + \dots + 1$$

deduced from the binomial formula.

Let $f = f_{p-1}$, so that

$$1^{p-1} + 2^{p-1} + \dots + n^{p-1} = f(n) = \frac{n^p}{p} + \dots$$

Choose the smallest integer $M \geq 1$ such that Mf has integer coefficients. Then $p \mid M$, since the leading coefficient of f is $\frac{1}{p}$. We know that $pn+r \mid Mf(n)$ for infinitely many n . But $pn+r$ also divides $p^p(Mf(n) - Mf(-\frac{r}{p}))$, thus $pn+r$ divides $p^p Mf(-\frac{r}{p})$ for infinitely many n . This yields $Mf(-\frac{r}{p}) = 0$. Using example 3.64, we obtain the existence of a polynomial g with integer coefficients such that

$$Mf(n) = (pn+r)g(n)$$

for all n . Since $f(n)$ is an integer for all n , $p \mid M$ and $\gcd(p, pn+r) = 1$, we deduce that $p \mid g(n)$ for all n . However $\deg g = \deg f - 1 = p-1$, thus Lagrange's theorem yields $g \equiv 0 \pmod{p}$. But then

$$\frac{M}{p} \cdot f(X) = (pX+r) \cdot \frac{g(X)}{p}$$

has integer coefficients, contradicting the minimality of M . The result follows. \square

35. (China 2016) Let p be an odd prime and a_1, a_2, \dots, a_p be integers. Prove that the following two conditions are equivalent:

- a) There is a polynomial P of degree $\leq \frac{p-1}{2}$ such that $P(i) \equiv a_i \pmod{p}$ for all $1 \leq i \leq p$;
 b) For any $1 \leq d \leq \frac{p-1}{2}$

$$\sum_{i=1}^p (a_{i+d} - a_i)^2 \equiv 0 \pmod{p},$$

where indices are taken modulo p .

Proof. The fact that a) implies b) is fairly easy. Indeed, if $a_i \equiv P(i) \pmod{p}$ and $\deg P \leq \frac{p-1}{2}$, then considering $Q(X) = P(X+d) - P(X)$ we have $\deg Q \leq \frac{p-3}{2}$, thus $\deg(Q^2) < p-1$ and so (by corollary 5.77)

$$\sum_{i=1}^p Q(i)^2 \equiv 0 \pmod{p},$$

which is exactly the content of part b).

Let us turn now to the interesting implication. Note that since p is odd and $1+d, 2+d, \dots, p+d$ is a complete residue system modulo p , the congruence in part b) is equivalent to

$$\sum_{i=1}^p a_i^2 \equiv \sum_{i=1}^p a_i a_{i+d}$$

for all $1 \leq d \leq \frac{p-1}{2}$. Consider a polynomial P of degree $\leq p-1$ such that $P(i) \equiv a_i \pmod{p}$ for $1 \leq i \leq p$ (P is actually unique by Lagrange's theorem). It is not difficult to construct such a polynomial: choose integers b_i such that $b_i \prod_{j \neq i} (i-j) \equiv 1 \pmod{p}$ and set

$$P(X) = \sum_{i=1}^p a_i b_i \prod_{j \neq i} (X-j).$$

Consider now the polynomial

$$Q(X) = \sum_{i=1}^p P(i)P(X+i).$$

The hypothesis of the problem becomes $Q(d) \equiv Q(0) \pmod{p}$ for $1 \leq d \leq \frac{p-1}{2}$. For such d we also have (using that $1-d, 2-d, \dots, p-d$ is a complete residue system mod p)

$$Q(-d) = \sum_{i=1}^p P(i)P(i-d) \equiv \sum_{j=1}^p P(j+d)P(j) = Q(d) \equiv Q(0) \pmod{p},$$

thus the congruence $Q(X) - Q(0) \equiv 0 \pmod{p}$ has at least p solutions. Since $\deg Q \leq p-1$, Lagrange's theorem yields $Q \equiv Q(0) \pmod{p}$, i.e. all coefficients of the polynomial $Q - Q(0)$ are multiples of p .

Finally, write

$$P(X) \equiv \alpha X^r + \beta X^{r-1} + \dots \pmod{p},$$

with $\alpha \neq 0$ and $r \leq p-1$. Assume that $r > \frac{p-1}{2}$ and let $k = 2r - (p-1)$. Note that $k > 0$ and $k \leq r$. Since

$$Q(X) = \sum_{i=1}^p P(i)[\alpha(X+i)^r + \beta(X+i)^{r-1} + \dots],$$

the coefficient of X^k in $Q(X) - Q(0)$ is

$$\sum_{i=1}^p P(i) \left(\alpha \binom{r}{k} i^{r-k} + \beta \binom{r-1}{k} i^{r-k-1} + \dots \right)$$

and this is 0 mod p by the previous discussion. We deduce that

$$\alpha \binom{r}{k} \sum_{i=1}^p P(i) i^{r-k} + \beta \binom{r-1}{k} \sum_{i=1}^p P(i) i^{r-k-1} + \dots \equiv 0 \pmod{p}.$$

Note that since $\deg(P \cdot X^{r-k-j}) = 2r - k - j = p - 1 - j$, we have (by corollary 5.77)

$$\sum_{i=1}^p P(i)i^{r-k} \equiv -\alpha \pmod{p}, \quad \sum_{i=1}^p P(i)i^{r-k-1} \equiv 0 \pmod{p}, \dots$$

Thus the previous congruence becomes

$$\alpha^2 \binom{r}{k} \equiv 0 \pmod{p}.$$

This is certainly impossible, since α is not a multiple of p and $p > r \geq k$. □

36. (USAMO 1999) Let p be an odd prime and let a, b, c, d be integers not divisible by p such that

$$\left\{ \frac{ra}{p} \right\} + \left\{ \frac{rb}{p} \right\} + \left\{ \frac{rc}{p} \right\} + \left\{ \frac{rd}{p} \right\} = 2$$

for all integers r not divisible by p (where $\{x\}$ is the fractional part of x). Prove that at least two of the numbers $a+b, a+c, a+d, b+c, b+d, c+d$ are divisible by p .

Proof. This is a very difficult problem! Let $r(x) \in \{0, 1, \dots, p-1\}$ be the remainder of $x \bmod p$, so that the hypothesis of the problem becomes

$$r(an) + r(bn) + r(cn) + r(dn) = 2p$$

for any n relatively prime to p . Call such a 4-tuple (a, b, c, d) good. Clearly if (a, b, c, d) is good, then so is (ka, kb, kc, kd) for any k which is not a multiple of p . Note also that if (a, b, c, d) is good, then $a+b+c+d \equiv 0 \pmod{p}$, since $r(a) + r(b) + r(c) + r(d) = 2p \equiv 0 \pmod{p}$.

Let

$$Q(x) = \frac{2r(x) - r(2x)}{p},$$

in other words $Q(x) = 0$ if $1 \leq r(x) \leq (p-1)/2$ and $Q(x) = 1$ if $(p-1)/2 < r(x) < p$. It follows from the first paragraph that

$$Q(ka) + Q(kb) + Q(kc) + Q(kd) = 2$$

for all $1 \leq k < p$ and all good 4-tuples (a, b, c, d) .

Next, choose¹ a polynomial $P(X)$ with integer coefficients, of degree at most $p-2$, such that $P(x) \equiv Q(x) \pmod{p}$ for all x not divisible by p , and define $R(X) = P(X+1) - P(X)$. Then $R(x) \equiv 0 \pmod{p}$ for $x = 1, \dots, \frac{p-3}{2}, \frac{p+1}{2}, \dots, p-2$ and $R(\frac{p-1}{2})$ is not divisible by p . We deduce from Lagrange's theorem that the coefficient of X^{p-3} in R is not divisible by p and hence the coefficient of X^{p-2} in P is not divisible by p . Next, letting

$$S(X) = P(Xa) + P(Xb) + P(Xc) + P(Xd),$$

the congruence $S(x) \equiv 2 \pmod{p}$ has at least $p-1$ solutions (by the second paragraph and the choice of P) and since $\deg S \leq p-2$, we deduce from Lagrange's theorem that $S(X) \equiv 2 \pmod{p}$, thus the coefficient of X^{p-2} in S is divisible by p . Combining the previous observations yields

$$a^{p-2} + b^{p-2} + c^{p-2} + d^{p-2} \equiv 0 \pmod{p},$$

which can be also written (by Fermat's little theorem)

$$a^{-1} + b^{-1} + c^{-1} + d^{-1} \equiv 0 \pmod{p},$$

where we write x^{-1} for the inverse of x modulo p . Since $a+b+c+d \equiv 0 \pmod{p}$, it follows that

$$a^{-1} + b^{-1} + c^{-1} \equiv (a+b+c)^{-1} \pmod{p}$$

and multiplying by $abc(a+b+c)$ we easily obtain $(a+b)(b+c)(c+a) \equiv 0 \pmod{p}$. By symmetry, we may assume that $a+b \equiv 0 \pmod{p}$. Since $a+b+c+d \equiv 0 \pmod{p}$, we also have $c+d \equiv 0 \pmod{p}$ and the result follows. \square

¹The existence of such a polynomial follows easily from Lagrange's interpolation theorem; see the solution of the previous problem for the simple argument.

37. Let n be a positive integer such that $p = 4n + 1$ is a prime. Prove that $n^n \equiv 1 \pmod{p}$.

Proof. We have $4n \equiv -1 \pmod{p}$, thus $4^n \cdot n^n \equiv (-1)^n \pmod{p}$. It suffices therefore to show that $4^n \equiv (-1)^n \pmod{p}$. But

$$4^n = 2^{\frac{p-1}{2}} \equiv (-1)^{\frac{p^2-1}{8}} = (-1)^{2n^2+n} = (-1)^n \pmod{p},$$

as needed. □

38. Let p be an odd prime. Prove that the number of integers $n \in \{1, 2, \dots, p-2\}$ such that n and $n+1$ are both quadratic residues mod p is $\frac{p-(-1)^{\frac{p-1}{2}}}{4} - 1$.

Proof. Let N be the desired number and observe that

$$N = \sum_{n=1}^{p-2} \frac{1}{4} \left(1 + \left(\frac{n}{p} \right) \right) \cdot \left(1 + \left(\frac{n+1}{p} \right) \right),$$

since the term indexed by n in the sum is 1 if n and $n+1$ are both quadratic residues mod p , and 0 otherwise. Expanding the product and rearranging terms yields

$$N = \frac{p-2}{4} + \frac{1}{4} \sum_{n=1}^{p-2} \left(\frac{n}{p} \right) + \frac{1}{4} \sum_{n=1}^{p-2} \left(\frac{n+1}{p} \right) + \frac{1}{4} \sum_{n=1}^{p-2} \left(\frac{n(n+1)}{p} \right).$$

On the other hand, we clearly have

$$\sum_{n=1}^{p-2} \left(\frac{n}{p} \right) = - \left(\frac{-1}{p} \right) = -(-1)^{\frac{p-1}{2}}, \quad \sum_{n=1}^{p-2} \left(\frac{n+1}{p} \right) = -1,$$

while proposition 5.111 gives

$$\sum_{n=1}^{p-2} \left(\frac{n(n+1)}{p} \right) = \sum_{n=1}^{p-1} \left(\frac{n(n+1)}{p} \right) = -1.$$

The result follows by combining these relations. □

Remark 8.16. One can also establish this result directly, by carefully analyzing the solutions of the congruence $y^2 - x^2 \equiv 1 \pmod{p}$.

39. (Gazeta Matematică) Prove that for any $n \geq 1$ the number $3^n + 2$ does not have prime divisors of the form $24k + 13$.

Proof. Suppose that $p \equiv 13 \pmod{24}$ is a prime divisor of $3^n + 2$ for some $n \geq 1$. Since $p \equiv 1 \pmod{4}$ and $p \equiv 1 \pmod{3}$, we deduce from the quadratic reciprocity law that

$$\left(\frac{3}{p}\right) = \left(\frac{p}{3}\right) = 1.$$

However since $p \equiv 5 \pmod{8}$, we also compute that

$$\left(\frac{-2}{p}\right) = \left(\frac{-1}{p}\right) \left(\frac{2}{p}\right) = -1.$$

This is a contradiction since $3^n \equiv -2 \pmod{p}$, but the left-hand side is a quadratic residue modulo p while the right-hand side is not. \square

40. Prove that there are infinitely many primes $p \equiv -1 \pmod{5}$.

Proof. Let $n > 1$ and consider a prime divisor p of $N = 5(n!)^2 - 1$ such that p is not congruent to 1 mod 5. Such a prime exists since otherwise N would be congruent to 1 mod 5, which is certainly not the case. Since $p \mid N$, we have $5 \equiv (5n!)^2 \pmod{p}$, thus 5 is a quadratic residue mod p . The quadratic reciprocity law then implies that p is a quadratic residue mod 5, and since p is not congruent to 1 mod 5, we must have $p \equiv -1 \pmod{5}$. Since $p > n$, varying n we obtain infinitely many primes $p \equiv -1 \pmod{5}$. \square

41. Let $p = a^2 + b^2$ be an odd prime, with a, b positive integers and a odd. Prove that a is a quadratic residue mod p .

Proof. It suffices to prove that any prime factor q of a is a quadratic residue mod p . Note that $q \neq p$ and that $p \equiv b^2 \pmod{q}$, thus $\left(\frac{p}{q}\right) = 1$. Using the quadratic reciprocity law, we deduce that

$$\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}} = 1,$$

the last equality being a consequence of the fact that $p \equiv 1 \pmod{4}$ (since p is a sum of two squares). The result follows. \square

42. Let n be a positive integer and let a be a divisor of $36n^4 - 8n^2 + 1$, such that 5 does not divide a . Prove that the remainder of a when divided by 20 is 1 or 9.

Proof. It suffices to prove the same statement for each prime factor of a , thus we may assume that $a = p$ is a prime.

First, since $p \mid (6n^2 - 1)^2 + (2n)^2$ and p does not divide simultaneously $6n^2 - 1$ and $2n$, we deduce that $p \equiv 1 \pmod{4}$. It remains to prove that $p \equiv 1, 4 \pmod{5}$. Since $p \neq 5$, by the quadratic reciprocity law this is equivalent to showing that $\left(\frac{5}{p}\right) = 1$, i.e. that 5 is a quadratic residue mod p . But $p \mid (6n^2 + 1)^2 - 5 \cdot (2n)^2$ and p does not divide $2n$ (otherwise p would also divide $6n^2 + 1$, impossible), which makes it clear that 5 is a quadratic residue mod p . The result follows. \square

43. Are there positive integers x, y, z such that $8xy = x + y + z^2$?

Proof. Assume that such integers exist, then they also satisfy

$$(8x - 1)(8y - 1) = 8z^2 + 1.$$

If p is a prime divisor of $8x - 1$ or $8y - 1$, then $p \mid 8z^2 + 1$, thus $(4z)^2 \equiv -2 \pmod{p}$ and so -2 is a quadratic residue mod p . We deduce that $p \equiv 1, 3 \pmod{8}$. Since $3 \cdot 3 \equiv 1 \pmod{8}$, the product of a finite number of primes of the form $8k + 1$ or $8k + 3$ is congruent to 1 or 3 modulo 8. We deduce that $8x - 1 \equiv 1, 3 \pmod{8}$, which is absurd. Thus the equation has no solutions. \square

44. (Komal A 618) Prove that there are no integers x, y such that

$$x^3 - x + 9 = 5y^2.$$

Proof. Assume that x, y are such integers. Note that the left-hand side is odd and a multiple of 3, thus y must be odd and a multiple of 3, say $y = 3t$. Then $x^3 - x + 9 \equiv 5y^2 \equiv 5 \pmod{8}$, thus $x^3 - x + 4 \equiv 0 \pmod{8}$, which implies that x is even and then that $4 \mid x$, say $x = 4z$. Note that z must be odd, as $x \equiv 4 \pmod{8}$. The equation can be rewritten

$$4z(16z^2 - 1) = 9(5t^2 - 1).$$

Note that $5t^2 - 1$ is not a multiple of 3 (for any t), thus the highest power of 3 dividing the right-hand side is 9. Also, one of the numbers z and $16z^2 - 1$ must be a multiple of 9 and the other must be relatively prime to 3. If $p \neq 3$ is a prime factor of z or $16z^2 - 1$, then $p \mid 5t^2 - 1$, thus 5 is a quadratic residue mod p . The quadratic reciprocity law (note that $p \neq 2$ as z and $16z^2 - 1$ are odd) yields $p \equiv \pm 1 \pmod{5}$. Thus all prime factors different from 3 of z and $16z^2 - 1$ are $\pm 1 \pmod{5}$. Since z is either relatively prime to 3 or of the form $9u$ with u relatively prime to 3, we deduce that $z \equiv \pm 1 \pmod{5}$. It follows that $16z^2 - 1 \equiv 0 \pmod{5}$, which is impossible since $9(5t^2 - 1)$ is not a multiple of 5. Thus the equation has no solutions. \square

45. Let p be an odd prime divisor of $n^4 - n^3 + 2n^2 + n + 1$, for some $n > 1$. Prove that $p \equiv 1, 4 \pmod{15}$.

Proof. Let

$$f(n) = 4(n^4 - n^3 + 2n^2 + n + 1).$$

One can directly check the equalities

$$f(n) = (2n^2 - n + 1)^2 + 3(n + 1)^2 = (2n^2 - n + 3)^2 - 5(n - 1)^2.$$

Note that $2n^2 - n + 1$ cannot be a multiple of 3 and $2n^2 - n + 3$ cannot be a multiple of 5. Thus $p \neq 3, 5$. We have

$$(2n^2 - n + 1)^2 \equiv -3(n + 1)^2 \pmod{p}$$

and p does not divide 3 or $n + 1$ (if $p \mid n + 1$, the previous congruence yields $p \mid (2 + 1 + 1)^2$, impossible), we obtain $\left(\frac{-3}{p}\right) = 1$, which by the quadratic reciprocity law can be rewritten $\left(\frac{p}{3}\right) = 1$. Thus $p \equiv 1 \pmod{3}$. Similarly, the congruence

$$(2n^2 - n + 3)^2 \equiv 5(n - 1)^2 \pmod{p}$$

combined with the quadratic reciprocity law yields $\left(\frac{p}{5}\right) = 1$, thus $p \equiv 1, 4 \pmod{5}$. Combining these we conclude that $p \equiv 1, 4 \pmod{15}$. \square

46. Prove that infinitely many primes don't divide any of the numbers $2^{n^2+1} - 3^n$ with $n \geq 1$.

Proof. Suppose that $p > 3$ is a prime that divides $2^{n^2+1} - 3^n$ for some n . Thus $2^{n^2+1} \equiv 3^n \pmod{p}$. Note that n and $n^2 + 1$ have different parity. We deduce that $\left(\frac{2}{p}\right) = 1$ or $\left(\frac{3}{p}\right) = 1$. The first case happens if and only if $p \equiv \pm 1 \pmod{8}$, while the second case happens (using the quadratic reciprocity law) if and only if $p \equiv \pm 1 \pmod{12}$. We conclude that p must be congruent to one of the numbers 1, 7, 11, 13, 17, 23 modulo 24. By Dirichlet's theorem we can find infinitely many primes congruent to 5 modulo 24, and the previous argument shows that none of them divides a number of the form $2^{n^2+1} - 3^n$. \square

47. a) (Gauss) Prove that an odd prime p can be written $a^2 + 2b^2$ for some integers a, b if and only if $p \equiv 1, 3 \pmod{8}$.
 b) (Euler, Lagrange) Prove that a prime $p \neq 3$ can be written $a^2 + 3b^2$ if and only if $p \equiv 1 \pmod{3}$.

Proof. a) Suppose first that $p = a^2 + 2b^2$. Then $a^2 \equiv -2b^2 \pmod{p}$. Since b is not a multiple of p , we deduce that $\left(\frac{-2}{p}\right) = 1$, which is equivalent to $p \equiv 1, 3 \pmod{8}$.

Conversely, assume that $p \equiv 1, 3 \pmod{8}$, so that by the above discussion -2 is a quadratic residue mod p . Let u be an integer such that

$u^2 \equiv -2 \pmod{p}$. Using Thue's lemma (see theorem 5.59), we can find integers a, b such that $|a|, |b| < \sqrt{p}$, a, b are not both 0 and $a \equiv ub \pmod{p}$. Then $a^2 \equiv u^2 b^2 \equiv -2b^2 \pmod{p}$, thus $p \mid a^2 + 2b^2$. Since $a^2 + 2b^2 < 3p$, we deduce that $p = a^2 + 2b^2$ or $2p = a^2 + 2b^2$. In the first case we are done, so assume that $2p = a^2 + 2b^2$, so that $a = 2c$ and $p = b^2 + 2c^2$. This finishes the proof.

b) The proof is very similar to that of a) and we leave the details to the reader. The key point is that -3 is a quadratic residue mod p if and only if $p \equiv 1 \pmod{3}$ (this follows directly from the quadratic reciprocity law). This immediately settles one implication. For the more difficult implication, one uses Thue's lemma as above to deduce the existence of integers a, b such that $|a|, |b| < \sqrt{p}$, a, b are not both 0 and $p \mid a^2 + 3b^2$. If $p = a^2 + 3b^2$ or $3p = a^2 + 3b^2$, one immediately obtains the desired conclusion. If $2p = a^2 + 3b^2$ one obtains a contradiction by taking the relation mod 3. \square

Remark 8.17. Proceeding in a similar (but slightly more technical) way, one can prove the following result of Lagrange and Gauss: if $p \neq 5$ is a prime, then p can be written $a^2 + 5b^2$ for some integers a, b if and only if $p \equiv 1, 9 \pmod{20}$, and $2p$ can be written $a^2 + 5b^2$ if and only if $p \equiv 3, 7 \pmod{20}$. In general, given an integer $n > 1$ it is a very delicate problem to describe all primes that can be represented $x^2 + ny^2$ for some integers x, y .

48. (Moldova TST 2005) Let $f, g : \mathbf{N} \rightarrow \mathbf{N}$ be functions with the properties:

- i) g is surjective;
- ii) $2f(n)^2 = n^2 + g(n)^2$ for all positive integers n .
- iii) $|f(n) - n| \leq 2004\sqrt{n}$ for all $n \in \mathbf{N}$.

Prove that f has infinitely many fixed points.

Proof. Let p_n be the sequence of prime numbers of the form $8k + 3$ (the fact that there are infinitely many such numbers follows from Dirichlet's theorem, or from example 5.131). Then $\left(\frac{2}{p_n}\right) = -1$ for all n . Letting x_n

be an integer such that $g(x_n) = p_n$, we obtain $2f(x_n)^2 = x_n^2 + p_n^2$, thus $2f(x_n)^2 \equiv x_n^2 \pmod{p_n}$. Since $\left(\frac{2}{p_n}\right) = -1$, we deduce that $p_n | x_n$ and $p_n | f(x_n)$. Thus there exist sequences of positive integers a_n, b_n such that $x_n = a_n p_n$ and $f(x_n) = b_n p_n$ for all n . Using ii) we obtain $2b_n^2 = a_n^2 + 1$ and iii) yields

$$\frac{2004}{\sqrt{x_n}} \geq \left| \frac{f(x_n)}{x_n} - 1 \right| = \left| \frac{b_n}{a_n} - 1 \right|.$$

Thus

$$\lim_{n \rightarrow \infty} \frac{\sqrt{a_n^2 + 1}}{a_n} = \sqrt{2}$$

and so $\lim_{n \rightarrow \infty} a_n = 1$. Therefore, starting from a certain n onwards, we have $a_n = 1 = b_n$, that is $f(p_n) = p_n$. The result follows. \square

49. (Romania TST 2004) Let p be an odd prime and let

$$f(x) = \sum_{i=1}^{p-1} \left(\frac{i}{p}\right) X^{i-1}.$$

a) Prove that f is divisible by $X - 1$ but not by $(X - 1)^2$ if and only if $p \equiv 3 \pmod{4}$;

b) Prove that if $p \equiv 5 \pmod{8}$ then f is divisible by $(X - 1)^2$ but not by $(X - 1)^3$.

Proof. a) Note that $f(1) = \sum_{i=1}^{p-1} \left(\frac{i}{p}\right) = 0$ and

$$\begin{aligned} f'(1) &= \sum_{i=1}^{p-1} (i-1) \left(\frac{i}{p}\right) = \sum_{i=1}^{p-1} i \left(\frac{i}{p}\right) = \sum_{i=1}^{p-1} (p-i) \left(\frac{p-i}{p}\right) \\ &= (-1)^{\frac{p-1}{2}} \sum_{i=1}^{p-1} (p-i) \left(\frac{i}{p}\right) = -(-1)^{\frac{p-1}{2}} f'(1). \end{aligned}$$

Hence for $p \equiv 1 \pmod{4}$ we have $f'(1) = 0$ and f is divisible by $(X-1)^2$. If $p \equiv 3 \pmod{4}$, then

$$f'(1) = \sum_{i=1}^{p-1} i \left(\frac{i}{p} \right) \equiv \sum_{i=1}^{p-1} i = \frac{p(p-1)}{2} \equiv 1 \pmod{2},$$

thus f is divisible by $X-1$ but not by $(X-1)^2$.

b) Using part a) we obtain

$$f''(1) = \sum_{i=1}^{p-1} (i^2 - 3i + 2) \left(\frac{i}{p} \right) = \sum_{i=1}^{p-1} i^2 \left(\frac{i}{p} \right)$$

and we need to show that this is nonzero. We will prove that $f''(1) \equiv 4 \pmod{8}$. Since $i^2 \left(\frac{2i}{p} \right) \equiv i \pmod{2}$ and $(2i-1)^2 \equiv 1 \pmod{8}$ for all i , we have

$$\begin{aligned} f''(1) &= \sum_{i=1}^{\frac{p-1}{2}} 4i^2 \left(\frac{2i}{p} \right) + \sum_{i=1}^{\frac{p-1}{2}} (2i-1)^2 \left(\frac{2i-1}{p} \right) \\ &\equiv 4 \sum_{i=1}^{\frac{p-1}{2}} i + \sum_{i=1}^{\frac{p-1}{2}} \left(\frac{2i-1}{p} \right) \equiv 4 + \sum_{i=1}^{\frac{p-1}{2}} \left(\frac{2i-1}{p} \right) \pmod{8}, \end{aligned}$$

since $\sum_{i=1}^{\frac{p-1}{2}} i = \frac{p^2-1}{8} \equiv 1 \pmod{2}$. It suffices therefore to prove the equality

$$\sum_{i=1}^{\frac{p-1}{2}} \left(\frac{2i-1}{p} \right) = 0.$$

For this, simply note that since $p \equiv 1 \pmod{4}$ we have

$$\sum_{i=1}^{\frac{p-1}{2}} \left(\frac{2i-1}{p} \right) = \sum_{i=1}^{\frac{p-1}{2}} \left(\frac{p-(2i-1)}{p} \right) = \sum_{i=1}^{\frac{p-1}{2}} \left(\frac{2i}{p} \right) = - \sum_{i=1}^{\frac{p-1}{2}} \left(\frac{2i-1}{p} \right)$$

since $\sum_{i=1}^{p-1} \left(\frac{i}{p} \right) = 0$. This finishes the proof. \square

50. For an odd prime p , let $f(p)$ be the number of solutions of the congruence $y^2 \equiv x^3 - x \pmod{p}$.

a) Prove that $f(p) = p$ for $p \equiv 3 \pmod{4}$.

b) Prove that if $p \equiv 1 \pmod{4}$ then

$$f(p) \equiv (-1)^{\frac{p+3}{4}} \left(\frac{\frac{p-1}{2}}{\frac{p-1}{4}} \right) \pmod{p}.$$

c) For which primes p do we have $f(p) = p$?

Proof. We have

$$f(p) = \sum_{x=0}^{p-1} \left(1 + \left(\frac{x^3 - x}{p} \right) \right) = p + \sum_{x=1}^{p-1} \left(\frac{x^3 - x}{p} \right).$$

a) If $p \equiv 3 \pmod{4}$, then for all x we have

$$\left(\frac{(-x)^3 - (-x)}{p} \right) = \left(\frac{-1}{p} \right) \cdot \left(\frac{x^3 - x}{p} \right) = - \left(\frac{x^3 - x}{p} \right),$$

thus

$$\sum_{x=1}^{p-1} \left(\frac{x^3 - x}{p} \right) = \sum_{x=1}^{\frac{p-1}{2}} \left(\frac{x^3 - x}{p} \right) + \sum_{x=1}^{\frac{p-1}{2}} \left(\frac{(-x)^3 - (-x)}{p} \right) = 0$$

and $f(p) = p$.

b) Writing $p - 1 = 4l$, we have (using the binomial formula and the congruence $\left(\frac{a}{p} \right) \equiv a^{\frac{p-1}{2}} \pmod{p}$)

$$\begin{aligned} \sum_{x=1}^{p-1} \left(\frac{x^3 - x}{p} \right) &\equiv \sum_{x=1}^{p-1} (x^3 - x)^{\frac{p-1}{2}} \equiv \sum_{x=1}^{p-1} \sum_{k=0}^{2l} \binom{2l}{k} x^{p-1+2l-2k} (-1)^k \\ &= \sum_{k=0}^{2l} (-1)^k \binom{2l}{k} \sum_{x=1}^{p-1} x^{p-1+2(l-k)} \pmod{p}. \end{aligned}$$

Note that

$$\sum_{x=1}^{p-1} x^{p-1+2(l-k)} \equiv \sum_{x=1}^{p-1} x^{2(l-k)} \pmod{p}$$

and (by corollary 5.77) the last sum is congruent to 0 mod p unless $l = k$ (as this is the only case when $2(l - k)$ is a multiple of $p - 1 = 4l$, for $0 \leq k \leq 2l$), in which case the sum is congruent to $-1 \pmod{p}$. We conclude that

$$\sum_{x=1}^{p-1} \left(\frac{x^3 - x}{p} \right) \equiv (-1)^{l+1} \binom{2l}{l} \pmod{p},$$

as desired.

c) Since $(-1)^{\frac{p+3}{4}} \binom{\frac{p-1}{2}}{\frac{p-1}{4}}$ is obviously not a multiple of p , we conclude that no $p \equiv 1 \pmod{4}$ is a solution of the problem. Combining this with part a) shows that the solutions are the primes of the form $4k + 3$. \square

51. Is there a polynomial f of degree 5 with integer coefficients such that f has no rational root and the congruence $f(x) \equiv 0 \pmod{p}$ has solutions for any prime p ?

Proof. The answer is positive, we will show that $f(X) = (X^2+3)(X^3+2)$ is a solution of the problem. Clearly f has degree 5 and no rational root. If $p = 2$, then $f(0) \equiv 0 \pmod{p}$, so assume that $p > 2$. If $p \equiv 1 \pmod{3}$, then a simple calculation using the quadratic reciprocity law shows that $\left(\frac{-3}{p}\right) = 1$ and so there is x such that $x^2 + 3 \equiv 0 \pmod{p}$. Thus the congruence $f(x) \equiv 0 \pmod{p}$ is solvable in this case. Suppose now that $p \equiv 2 \pmod{3}$, then the map $x \mapsto x^3 \pmod{p}$ is bijective (see theorem 5.29), so the congruence $x^3 + 2 \equiv 0 \pmod{p}$ is solvable. Thus f is a solution of the problem. \square

Remark 8.18. For more details on this problem, the reader can consult the article "Polynomials $(x^3 - n)(x^2 + 3)$ solvable modulo any integer" by A. M. Hyde, P. D. Lee and B. K. Spearman, published in the American Mathematical Monthly, vol. 121, no. 4, p. 355-358.

52. Let p be an odd prime and let a be an integer not divisible by p . Let $N(a)$ be the number of solutions of the congruence $y^2 \equiv x^3 + ax \pmod{p}$ and let

$$S(a) = \sum_{k=0}^{p-1} \left(\frac{k^3 + ak}{p} \right).$$

- 1) Prove that $N(a) = p + S(a)$.
- 2) Prove that if $p \equiv 3 \pmod{4}$ then $S(a) = 0$ for all a , hence $N(a) = p$. We assume from now on that $p \equiv 1 \pmod{4}$.
- 3) Prove that if b is not a multiple of p , then

$$S(ab^2) = \left(\frac{b}{p} \right) S(a).$$

- 4) Prove that

$$\sum_{a=0}^{p-1} S(a)^2 = 2p(p-1)$$

and that if $A = S(-1)$ and $B = S(a)$ for any quadratic non-residue a , then

$$A^2 + B^2 = 4p.$$

- 5) Prove that $A \equiv -(p+1) \pmod{8}$.
- 6) Deduce the following theorem of Jacobsthal: let $p \equiv 1 \pmod{4}$ be a prime and write $p = a^2 + b^2$ with a, b integers, a odd and $a \equiv -\frac{p+1}{2} \pmod{4}$. Then the congruence $y^2 \equiv x^3 - x \pmod{p}$ has $p+2a$ solutions.

Proof. 1) This is clear, since for each x the congruence $y^2 \equiv x^3 + ax \pmod{p}$ has $1 + \left(\frac{x^3 + ax}{p} \right)$ solutions.

- 2) Since

$$(p-k)^3 + a(p-k) \equiv -k^3 - ak = -(k^3 + ak) \pmod{p}$$

and $p \equiv 3 \pmod{4}$, we obtain for all k

$$\left(\frac{(p-k)^3 + a(p-k)}{p} \right) = \left(\frac{-1}{p} \right) \cdot \left(\frac{k^3 + ak}{p} \right) = - \left(\frac{k^3 + ak}{p} \right).$$

Adding these relations for $k = 0, 1, \dots, \frac{p-1}{2}$ gives $S(a) = 0$ and $N(a) = p$.

3) Since the remainders of $0, b, 2b, \dots, (p-1)b$ when divided by p are a permutation of $0, 1, \dots, p-1$, we obtain

$$\begin{aligned} S(ab^2) &= \sum_{k=0}^{p-1} \left(\frac{k^3 + ab^2k}{p} \right) = \sum_{k=0}^{p-1} \left(\frac{(kb)^3 + ab^2(kb)}{p} \right) = \sum_{k=0}^{p-1} \left(\frac{b^3(k^3 + ak)}{p} \right) \\ &= \left(\frac{b}{p} \right) \sum_{k=0}^{p-1} \left(\frac{k^3 + ak}{p} \right) = \left(\frac{b}{p} \right) S(a). \end{aligned}$$

4) We have

$$\begin{aligned} \sum_{a=0}^{p-1} S(a)^2 &= \sum_{a=0}^{p-1} \sum_{k,l=0}^{p-1} \left(\frac{k}{p} \right) \cdot \left(\frac{k^2 + a}{p} \right) \cdot \left(\frac{l}{p} \right) \cdot \left(\frac{l^2 + b}{p} \right) \\ &= \sum_{k,l=0}^{p-1} \left(\frac{kl}{p} \right) \sum_{a=0}^{p-1} \left(\frac{(k^2 + a)(l^2 + a)}{p} \right). \end{aligned}$$

For fixed k, l , the inner sum $\sum_{a=0}^{p-1} \left(\frac{(k^2+a)(l^2+a)}{p} \right)$ equals -1 when k^2 and l^2 are not congruent modulo p and $p-1$ otherwise, by proposition 5.111. It follows that

$$\sum_{a=0}^{p-1} S(a)^2 = p \sum_{0 \leq k, l \leq p-1, k^2 \not\equiv l^2 \pmod{p}} \left(\frac{kl}{p} \right) - \sum_{k,l=0}^{p-1} \left(\frac{kl}{p} \right).$$

Next, we have

$$\sum_{k,l=0}^{p-1} \left(\frac{kl}{p} \right) = \left(\sum_{k=0}^{p-1} \left(\frac{k}{p} \right) \right)^2 = 0.$$

Finally, note that if $k^2 \equiv l^2 \pmod{p}$ and k, l are not multiples of p , then $\left(\frac{kl}{p}\right) = 1$, since $k \equiv \pm l \pmod{p}$ and $\left(\frac{-1}{p}\right) = 1$. It follows that

$$\sum_{0 \leq k, l \leq p-1, k^2 \equiv l^2 \pmod{p}} \left(\frac{kl}{p}\right) = 2(p-1)$$

and so

$$\sum_{a=0}^{p-1} S(a)^2 = 2p(p-1).$$

The second part follows, since $S(a) = A$ whenever $\left(\frac{a}{p}\right) = 1$ and $S(a) = B$ whenever $\left(\frac{a}{p}\right) = -1$, thus

$$\sum_{a=0}^{p-1} S(a)^2 = S(0)^2 + (p-1)(A^2 + B^2) = (p-1)(A^2 + B^2).$$

5) This is fairly tricky. Note that

$$A = \sum_{k=0}^{p-1} \left(\frac{k^3 - k}{p}\right) = \sum_{k=0}^{p-1} \left(\frac{k-1}{p}\right) \cdot \left(\frac{k}{p}\right) \cdot \left(\frac{k+1}{p}\right).$$

Next, consider the expression

$$E = \sum_{k=0}^{p-1} \left(1 + \left(\frac{k-1}{p}\right)\right) \cdot \left(1 + \left(\frac{k}{p}\right)\right) \cdot \left(1 + \left(\frac{k+1}{p}\right)\right),$$

and note that for $k \neq 0, 1, p-2$ the corresponding term in the sum is the product of three even numbers, thus a multiple of 8. When $k = 0, 1, p-1$ we use the fact that $\left(\frac{-1}{p}\right) = 1$ to obtain

$$E \equiv 4 + 4\left(1 + \left(\frac{2}{p}\right)\right) \equiv 4 \pmod{8}.$$

On the other hand, brutally expanding we obtain (using proposition 5.111)

$$\begin{aligned} E &= p + \sum_{j \in \{-1, 0, 1\}} \sum_{k=0}^{p-1} \left(\frac{k+j}{p} \right) + \sum_{i < j \in \{-1, 0, 1\}} \sum_{k=0}^{p-1} \left(\frac{(k+i)(k+j)}{p} \right) + A \\ &= p + \sum_{j \in \{-1, 0, 1\}} 0 + \sum_{i < j \in \{-1, 0, 1\}} (-1) + A = p - 3 + A. \end{aligned}$$

Thus $p - 3 + A \equiv 4 \pmod{p}$, which yields $A \equiv -(p+1) \pmod{8}$.

6) It is immediate to check that A and B are even numbers, hence

$$p = \left(\frac{A}{2} \right)^2 + \left(\frac{B}{2} \right)^2$$

is the unique way to write p as the sum of two squares, up to sign. The previous part determines uniquely A and the result follows easily. \square

53. (Mathematical Reflections) Find all primes p with the following property: whenever a, b, c are integers and $p \mid a^2b^2 + b^2c^2 + c^2a^2 + 1$, we also have $p \mid a^2b^2c^2(a^2 + b^2 + c^2 + a^2b^2c^2)$.

Proof. This is a very difficult problem! The answer is 2, 3, 5, 13 and 17. Define $X_1(p)$ as the set of solutions in $\{0, 1, \dots, p-1\}^3$ of the congruence $a^2b^2 + b^2c^2 + c^2a^2 + 1 \equiv 0 \pmod{p}$, and similarly let $X_2(p)$ be the set of solutions of the congruence $a^2b^2c^2(a^2 + b^2 + c^2 + a^2b^2c^2) \equiv 0 \pmod{p}$. We want to find all primes p for which $X_1(p) \subset X_2(p)$.

First, we prove that 2, 3, 5, 13 and 17 are solutions of the problem. Suppose that $(a, b, c) \in X_1(p) \setminus X_2(p)$ for some prime p . Letting $x = a^2$, $y = b^2$, $z = c^2$, it follows that x, y, z are quadratic residues modulo p , $xy + yz + zx + 1 \equiv 0 \pmod{p}$ and $xyz(x + y + z + xyz)$ is not divisible by p . Since

$$\varepsilon(xy + yz + zx + 1) + (x + y + z + xyz) = (x + \varepsilon)(y + \varepsilon)(z + \varepsilon)$$

for $\varepsilon = \pm 1$, we deduce that x, y, z are not equal to 0 or ± 1 modulo p . This already excludes the cases $p = 2$, $p = 3$ and $p = 5$. Moreover, we cannot

have $x + y \equiv 0 \pmod{p}$ (and similar congruences obtained by permuting the variables x, y, z), as otherwise we would also have $xy + 1 \equiv 0 \pmod{p}$ and then $x + y + z + xyz \equiv 0 \pmod{p}$, a contradiction. Similarly, we cannot have $xy + 1 \equiv 0 \pmod{p}$ (and similarly with permutations of x, y, z). Finally, x, y, z must be pairwise distinct if $p \in \{13, 17\}$, for one tediously checks that for these primes the congruence $x^2 + 2zx + 1 \equiv 0 \pmod{p}$ has no solutions with x, z quadratic residues modulo p different from $0, \pm 1$. Using these observations and the fact that the quadratic residues modulo 13 (respectively 17) are $0, \pm 1, \pm 3, \pm 4$ (respectively $0, \pm 1, \pm 2, \pm 4, \pm 8$), one easily (but tediously!) checks that there are no triples (x, y, z) with all the previous properties. It follows that 2, 3, 5, 13 and 17 are solutions of the problem.

Next, we prove that if $p > 3$ is of the form $4k + 3$, then $X_1(p)$ is nonempty and disjoint from $X_2(p)$, hence p is not a solution of the problem. Pick an integer c such that c^2 is not congruent to 0 or 1 mod p (such c exists, since $p > 3$). Note that if $a \in \{0, 1, \dots, \frac{p-1}{2}\}$, then p does not divide $a^2 + c^2$ (as $p \equiv 3 \pmod{4}$ and p does not divide c), thus we can define a map f by imposing $f : \{0, 1, \dots, \frac{p-1}{2}\} \rightarrow \{0, 1, \dots, p-1\}$,

$$f(a)(a^2 + c^2) \equiv -(a^2c^2 + 1) \pmod{p}.$$

We claim that this map is injective. Indeed, if $f(a) = f(a_1)$, then an easy computation gives $(a^2 - a_1^2)(c^4 - 1) \equiv 0 \pmod{p}$, hence $a \equiv a_1 \pmod{p}$ (because $c^2 \pm 1$ are not divisible by p). Since f is injective and since there are $\frac{p+1}{2}$ quadratic residues mod p , it follows that there are $a, b \in \{0, 1, \dots, p-1\}$ such that $f(a) \equiv b^2 \pmod{p}$, which is equivalent to $(a, b, c) \in X_1(p)$. Hence $X_1(p) \neq \emptyset$.

Next, suppose that $(a, b, c) \in X_1(p) \cap X_2(p)$. Since $p \equiv 3 \pmod{4}$, p does not divide abc , hence $a^2(b^2c^2 + 1) + b^2 + c^2 \equiv 0 \pmod{p}$ and $a^2(b^2 + c^2) + b^2c^2 + 1 \equiv 0 \pmod{p}$. This yields $(a^4 - 1)(b^2 + c^2) \equiv 0 \pmod{p}$, then $a^2 \equiv 1 \pmod{p}$ and finally $(1 + b^2)(1 + c^2) \equiv 0 \pmod{p}$, a contradiction with $p \equiv 3 \pmod{4}$.

Finally, suppose that $p \equiv 1 \pmod{4}$ and $p > 17$. We will construct an element of $X_1(p)$ which is not in $X_2(p)$, finishing the solution. Since

$p \equiv 1 \pmod{4}$, there exists $x \in \mathbf{Z}$ such that $x^2 + 1 \equiv 0 \pmod{p}$. We will need the following

Lemma 8.19. *The congruence $a^2 + ab + b^2 \equiv x \pmod{p}$ has at least $p - 1$ solutions.*

Proof. Write the congruence as $(2a + b)^2 + 3b^2 \equiv 4x \pmod{p}$. So it is enough to prove that the congruence $u^2 + 3v^2 \equiv t \pmod{p}$ has at least $p - 1$ solutions when p does not divide t . But the number of solutions is $p + \sum_{v=0}^{p-1} \left(\frac{t-3v^2}{p} \right)$ and the result follows from proposition 5.111. \square

Now let S be the set of solutions of the previous congruence. For each $(a, b) \in S$ we have an element (a, b, c) of $X_1(p)$, where $c = -a - b$. Indeed,

$$\begin{aligned} a^2b^2 + b^2c^2 + c^2a^2 + 1 &\equiv (ab + bc + ca)^2 + 1 \equiv (a^2 + ab + b^2)^2 + 1 \\ &\equiv x^2 + 1 \equiv 0 \pmod{p}, \end{aligned}$$

hence $(a, b, c) \in X_1(p)$. Now we bound the number of these elements that lie in $X_2(p)$. Suppose that $(a, b, c) \in X_2(p)$. If $a \equiv 0 \pmod{p}$, then $b^2 \equiv x \pmod{p}$ and (a, b) takes at most two values mod p . Similarly the cases $b \equiv 0 \pmod{p}$ and $c \equiv 0 \pmod{p}$ yield each at most 2 values for $(a, b) \pmod{p}$, hence we have at most 6 elements of this form in $X_2(p)$. The other possibility is that $a^2 + b^2 + c^2 + a^2b^2c^2 \equiv 0 \pmod{p}$. Since

$$a^2 + b^2 + c^2 \equiv 2(a^2 + b^2 + ab) \equiv 2x \pmod{p}$$

and

$$a^2c^2 \equiv (a(a + b))^2 \equiv (x - b^2)^2 \pmod{p},$$

we obtain $2x + b^2(x - b^2)^2 \equiv 0 \pmod{p}$. This congruence has at most 6 solutions by Lagrange's theorem, and each solution b corresponds to at most two pairs (a, b) . Hence $X_2(p)$ contains at most 12 elements of this type. Thus in total $X_2(p)$ contains at most 18 of these elements of $X_1(p)$. Since $p > 17$ and $p \equiv 1 \pmod{4}$, we have at least one element of $X_1(p)$ which is not in $X_2(p)$. \square

54. Let n be a positive integer and let $p \geq 2n + 1$ be a prime. Prove that

$$\binom{2n}{n} \equiv (-4)^n \binom{\frac{p-1}{2}}{n} \pmod{p}.$$

Proof. We have

$$\begin{aligned} \binom{2n}{n} &= \frac{(2n)!}{n!^2} = \frac{2 \cdot 4 \cdot \dots \cdot (2n)}{n!} \cdot \frac{1 \cdot 3 \cdot \dots \cdot (2n-1)}{n!} \\ &= 2^n \cdot \frac{1 \cdot 3 \cdot \dots \cdot (2n-1)}{n!}. \end{aligned}$$

Since $n < p$ we have $\gcd(n!, p) = 1$ and so the desired congruence is equivalent (after multiplication by $n!$ and division by 2^n) with

$$1 \cdot 3 \cdot \dots \cdot (2n-1) \equiv (-2)^n \cdot \frac{p-1}{2} \left(\frac{p-1}{2} - 1 \right) \cdot \dots \cdot \left(\frac{p-1}{2} - n + 1 \right) \pmod{p}.$$

This congruence follows by multiplying the congruences

$$(-2) \left(\frac{p-1}{2} - j \right) \equiv 2j + 1$$

for $0 \leq j \leq n-1$. □

55. (Mathematical Reflections O 96) Prove that if $q \geq p$ are primes, then

$$pq \mid \binom{p+q}{p} - \binom{q}{p} - 1.$$

Proof. If $p = q$, this comes down to $\binom{2p}{p} \equiv 2 \pmod{p^2}$ and has already been proved (see example 5.157). So suppose that $q > p$. By Vandermonde's identity

$$\binom{p+q}{p} = \binom{p}{p} \binom{q}{0} + \binom{p}{p-1} \binom{q}{1} + \dots + \binom{p}{0} \binom{q}{p}$$

and each term in the sum except for the first and last one is a multiple of pq . □

56. (Hewgill) Let $n = n_0 + 2n_1 + \dots + 2^d n_d$ be the binary representation of an integer $n > 1$ and let S be the subset of $\{0, 1, \dots, n\}$ consisting of those k such that $\binom{n}{k}$ is odd. Prove that

$$\sum_{k \in S} 2^k = F_0^{n_0} F_1^{n_1} \dots F_d^{n_d},$$

where $F_k = 2^{2^k} + 1$ is the k th Fermat number.

Proof. By Lucas' theorem, the elements of S are precisely those

$$k = k_0 + 2k_1 + \dots + 2^d k_d \in \{0, 1, \dots, n\}$$

with $0 \leq k_i \leq n_i$ for all $0 \leq i \leq d$. We deduce that

$$\sum_{k \in S} 2^k = \sum_{k_0=0}^{n_0} \sum_{k_1=0}^{n_1} \dots \sum_{k_d=0}^{n_d} 2^{k_0} \cdot 2^{2k_1} \cdot \dots \cdot 2^{2^d k_d} = \left(\sum_{k_0=0}^{n_0} 2^{k_0} \right) \cdot \dots \cdot \left(\sum_{k_d=0}^{n_d} 2^{2^d k_d} \right).$$

It suffices to observe that since $n_0, \dots, n_d \in \{0, 1\}$, we have

$$\sum_{k_0=0}^{n_0} 2^k = (2+1)^{n_0}, \dots, \sum_{k_d=0}^{n_d} 2^{2^d k_d} = (2^{2^d} + 1)^{n_d}. \quad \square$$

57. (Calkin) Let a be a positive integer and let

$$x_n = \sum_{k=0}^n \binom{n}{k}^a$$

for $n \geq 1$. Let p be a prime, n an integer greater than 1 and let

$$n = n_0 + pn_1 + \dots + p^d n_d$$

be its base p representation. Prove that

$$x_n \equiv \prod_{i=0}^d x_{n_i} \pmod{p}.$$

Proof. Clearly

$$x_n \equiv \sum_{k \in S_n} \binom{n}{k}^a \pmod{p},$$

where S_n is the set of those $k \in \{0, 1, \dots, n\}$ for which p does not divide $\binom{n}{k}$. By Lucas' theorem the set S_n consists of the numbers

$$k = k_0 + pk_1 + \dots + p^d k_d$$

with $0 \leq k_i \leq n_i$ for all $0 \leq i \leq d$, and moreover for each $k \in S$ we have

$$\binom{n}{k} \equiv \prod_{i=0}^d \binom{n_i}{k_i} \pmod{p}.$$

We obtain therefore

$$x_n \equiv \sum_{k_0=0}^{n_0} \sum_{k_1=0}^{n_1} \dots \sum_{k_d=0}^{n_d} \prod_{i=0}^d \binom{n_i}{k_i}^a = \prod_{i=0}^d \left(\sum_{k_i=0}^{n_i} \binom{n_i}{k_i}^a \right) = \prod_{i=0}^d x_{n_i} \pmod{p},$$

as needed. \square

58. Let p be a prime and let k be an odd integer such that $p-1$ does not divide $k+1$. Prove that

$$\sum_{j=1}^{p-1} \frac{1}{j^k} \equiv 0 \pmod{p^2}.$$

Proof. It suffices to prove that

$$2 \sum_{j=1}^{p-1} \frac{1}{j^k} \equiv 0 \pmod{p^2}.$$

But

$$2 \sum_{j=1}^{p-1} \frac{1}{j^k} = \sum_{j=1}^{p-1} \left(\frac{1}{j^k} + \frac{1}{(p-j)^k} \right) = \sum_{j=1}^{p-1} \frac{j^k + (p-j)^k}{j^k(p-j)^k}$$

and since k is odd, the binomial formula gives

$$j^k + (p-j)^k \equiv kpj^{k-1} \pmod{p^2}.$$

It suffices therefore to prove that

$$\sum_{j=1}^{p-1} \frac{1}{j(p-j)^k} \equiv 0 \pmod{p}$$

or equivalently that

$$\sum_{j=1}^{p-1} \frac{1}{j^{k+1}} \equiv 0 \pmod{p}.$$

This follows from proposition 5.149. □

59. (Tuymaada 2012) Let $p = 4k + 3$ be a prime and write

$$\frac{1}{0^2 + 1} + \frac{1}{1^2 + 1} + \cdots + \frac{1}{(p-1)^2 + 1} = \frac{m}{n}$$

for some relatively prime numbers m, n . Prove that $p \mid 2m - n$.

Proof. Since $p \equiv 3 \pmod{4}$, the numbers $0^2 + 1, 1^2 + 1, \dots, (p-1)^2 + 1$ are not multiples of p by corollary 5.28. Next, we argue as in the proof of Wilson's theorem, by creating pairs of the form (x, y) with $x, y \in \{2, 3, \dots, p-2\}$ uniquely determined by $xy \equiv 1 \pmod{p}$. Note that for such a pair (x, y) ,

$$\frac{1}{x^2 + 1} + \frac{1}{y^2 + 1} \equiv \frac{1}{x^2 + 1} + \frac{1}{1 + \frac{1}{x^2}} \equiv 1 \pmod{p}.$$

Since the congruence $x^2 \equiv 1 \pmod{p}$ has no solutions in $\{2, \dots, p-2\}$ we deduce that

$$\frac{1}{2^2 + 1} + \cdots + \frac{1}{(p-2)^2 + 1} \equiv \frac{p-3}{2} \pmod{p},$$

hence

$$\frac{1}{0^2+1} + \frac{1}{1^2+1} + \dots + \frac{1}{(p-1)^2+1} \equiv \frac{p-3}{2} + 1 + 2 \cdot \frac{1}{2} = \frac{p+1}{2} \pmod{p}$$

The result follows. \square

60. (IMO Shortlist 2012) Find all integers $m \geq 2$ such that $n \mid \binom{n}{m-2n}$ for any integer $n \in [\frac{m}{3}, \frac{m}{2}]$.

Proof. We will prove that the solutions of the problem are exactly the prime numbers. If m is a prime number, then for any $n \in [\frac{m}{3}, \frac{m}{2}]$ the number

$$(m-2n) \binom{n}{m-2n} = n \binom{n-1}{m-2n-1}$$

is a multiple of n and since $\gcd(n, m-2n) = 1$ we obtain $n \mid \binom{n}{m-2n}$.

Conversely, let m be a solution of the problem. If m is even, choosing $n = \frac{m}{2}$ yields $\frac{m}{2} \mid 1$ and so $m = 2$. Assume that m is odd and let us suppose that m is composite. Then the smallest prime factor p of m is less than or equal to $\frac{m}{3}$. Let $n = \frac{m-p}{2}$, so $n \in [\frac{m}{3}, \frac{m}{2}]$ and by assumption $n \mid \binom{n}{p}$. We obtain $p \mid (n-1)(n-2)\dots(n-p+1)$, impossible since $p \mid n$. Thus m must be a prime number. \square

61. (Putnam 1991) Prove that for all odd primes p we have

$$\sum_{k=0}^p \binom{p}{k} \binom{p+k}{k} \equiv 2^p + 1 \pmod{p^2}.$$

Proof. For $1 \leq k \leq p-1$ we have

$$\binom{p+k}{k} = \frac{(p+1)(p+2)\dots(p+k)}{k!} \equiv 1 \pmod{p}$$

and $\binom{p}{k} \equiv 0 \pmod{p}$, hence $\binom{p}{k} \binom{p+k}{k} \equiv \binom{p}{k} \pmod{p^2}$. Thus the congruence is equivalent to

$$1 + \binom{2p}{p} + \sum_{k=1}^{p-1} \binom{p}{k} \equiv 2^p + 1 \pmod{p^2}$$

and then (using the binomial formula) to $\binom{2p}{p} \equiv 2 \pmod{p^2}$. This has already been established in example 5.157. \square

62. (ELMO Shortlist 2011) Prove that if p is a prime greater than 3 then

$$\sum_{k=0}^{\frac{p-1}{2}} \binom{p}{k} 3^k \equiv 2^p - 1 \pmod{p^2}.$$

Proof. We have for $1 \leq k \leq \frac{p-1}{2}$

$$\binom{p}{k} \equiv \frac{p}{k} (-1)^{k-1} \equiv 2(-1)^k \frac{p}{2k} (-1)^{2k-1} \equiv 2(-1)^k \binom{p}{2k} \pmod{p^2}.$$

The congruence is thus equivalent to

$$1 + 2 \sum_{k=1}^{\frac{p-1}{2}} \binom{p}{2k} (-3)^k + 1 \equiv 2^p - 1 \pmod{p^2}$$

or

$$2 \sum_{k=0}^{\frac{p-1}{2}} \binom{p}{2k} (-3)^k \equiv 2^p \pmod{p^2}.$$

Let $\alpha = i\sqrt{3} \in \mathbf{C}$, then $-3 = \alpha^2$ and the binomial formula yields

$$2 \sum_{k=0}^{\frac{p-1}{2}} \binom{p}{2k} \alpha^{2k} = (1 + \alpha)^p + (1 - \alpha)^p.$$

The result follows now from the equality $(1 + \alpha)^p + (1 - \alpha)^p = 2^p$, which itself follows from $1 + \alpha = 2e^{\frac{i\pi}{3}}$, $1 - \alpha = 2e^{-\frac{i\pi}{3}}$ and $\cos(\frac{\pi}{3}) = \frac{1}{2}$. \square

63. (IberoAmerican Olympiad 2005) Let $p > 3$ be a prime. Prove that

$$\sum_{i=1}^{p-1} \frac{1}{i^p} \equiv 0 \pmod{p^3}.$$

Proof. Let S denote the left-hand side. Then

$$2S = \sum_{i=1}^{p-1} \left(\frac{1}{i^p} + \frac{1}{(p-i)^p} \right) = \sum_{i=1}^{p-1} \frac{i^p + (p-i)^p}{i^p(p-i)^p}.$$

Using the binomial formula we obtain

$$i^p + (p-i)^p \equiv p^2(-i)^{p-1} = p^2 i^{p-1} \pmod{p^3},$$

thus it suffices to show that

$$\sum_{i=1}^{p-1} \frac{1}{i(p-i)^p} \equiv 0 \pmod{p}.$$

Since $(p-i)^p \equiv p-i \equiv -i \pmod{p}$, we are further reduced to showing that

$$\sum_{i=1}^{p-1} \frac{1}{i^2} \equiv 0 \pmod{p},$$

which has already been established in proposition 5.149. \square

64. (AMM) Let $C_n = \frac{1}{n+1} \binom{2n}{n}$ be the n th Catalan number. Prove that

$$C_1 + C_2 + \dots + C_n \equiv 1 \pmod{3}$$

if and only if $n+1$ has at least one digit equal to 2 in base 3.

Proof. One easily checks the equality

$$\binom{2k+2}{k+1} - 4\binom{2k}{k} = -2C_k,$$

which shows that

$$C_k \equiv \binom{2k+2}{k+1} - \binom{2k}{k} \pmod{3}.$$

It follows that

$$C_1 + C_2 + \dots + C_n \equiv \binom{2n+2}{n+1} - \binom{2}{1} = 1 + \binom{2n+2}{n+1} \pmod{3}.$$

Thus we need to prove that $3 \mid \binom{2n+2}{n+1}$ if and only if $n+1$ has at least one digit equal to 2 in base 3. This follows directly from example 5.146. \square

65. Prove that for any prime $p > 5$ we have

$$\left(1 + p \sum_{k=1}^{p-1} \frac{1}{k}\right)^2 \equiv 1 - p^2 \sum_{k=1}^{p-1} \frac{1}{k^2} \pmod{p^5}.$$

Proof. Letting

$$x_j = \sum_{k=1}^{p-1} \frac{1}{k^j},$$

the congruence is easily seen to be equivalent to

$$2x_1 + p(x_1^2 + x_2) \equiv 0 \pmod{p^4}.$$

Since $x_1 \equiv 0 \pmod{p^2}$, this further reduces to $2x_1 + px_2 \equiv 0 \pmod{p^4}$. Note that

$$2x_1 = \sum_{k=1}^{p-1} \left(\frac{1}{k} + \frac{1}{p-k}\right) = -p \sum_{k=1}^{p-1} \frac{1}{k^2(1-\frac{p}{k})}.$$

Note² that if $z \equiv 0 \pmod{p}$ then

$$1 - (1-z)(1+z+z^2) = z^3 \equiv 0 \pmod{p^3},$$

²This is motivated by the identity $\frac{1}{1-z} = 1 + z + z^2 + \dots$

thus

$$\frac{1}{1-z} \equiv 1 + z + z^2 \pmod{p^3}.$$

Combining this with $x_3 \equiv 0 \pmod{p^2}$ (see example 58) and $x_4 \equiv 0 \pmod{p}$, we obtain

$$2x_1 \equiv -p \sum_{k=1}^{p-1} \frac{1}{k^2} \left(1 + \frac{p}{k} + \frac{p^2}{k^2} \right) = -px_2 - p^2x_3 - p^3x_4 \equiv -px_2 \pmod{p^4},$$

as desired. \square

66. (USA TST 2002) Let $p > 5$ be a prime number. For any integer x , define

$$f_p(x) = \sum_{k=1}^{p-1} \frac{1}{(px+k)^2}$$

Prove that $f_p(x) \equiv f_p(y) \pmod{p^3}$ for all positive integers x, y .

Proof. Observe that

$$\frac{1}{(px+k)^2} = \frac{1}{k^2} \cdot \frac{1}{\left(1 - \left(-\frac{px}{k}\right)\right)^2}.$$

The identity

$$\frac{1}{(1-z)^2} = 1 + 2z + 3z^2 + \dots$$

suggests the congruence

$$\frac{1}{(1-z)^2} \equiv 1 + 2z + 3z^2 \pmod{p^3}$$

whenever $z \equiv 0 \pmod{p}$. Once we guess that this is true, it is very easy to prove it: an explicit computation shows that

$$1 - (1-z)^2(1+2z+3z^2) = 4z^3 - 3z^4 \equiv 0 \pmod{p^3}.$$

We deduce that

$$f(x) \equiv \sum_{k=1}^{p-1} \frac{1}{k^2} \cdot \frac{1}{(1 - (-\frac{px}{k}))^2} \equiv \sum_{k=1}^{p-1} \frac{1}{k^2} \left(1 - \frac{2px}{k} + \frac{3p^2x^2}{k^2} \right) \pmod{p^3}.$$

Since

$$\sum_{k=1}^{p-1} \frac{1}{k^3} \equiv 0 \pmod{p^2} \quad \text{and} \quad \sum_{k=1}^{p-1} \frac{1}{k^4} \equiv 0 \pmod{p},$$

we conclude that

$$f_p(x) \equiv \sum_{k=1}^{p-1} \frac{1}{k^2} \equiv f_p(y) \pmod{p^3}$$

for all x, y , as desired. \square

8.5 p -adic valuations and the distribution of primes

1. (Russia 2000) Prove that there is a partition of \mathbf{N} with 100 sets such that if $a, b, c \in \mathbf{N}$ satisfy $a + 99b = c$, then at least two of the numbers a, b, c belong to the same set.

Proof. Let the i th class consist of those n for which $v_2(n) \equiv i \pmod{100}$ for $1 \leq i \leq 100$. If $a + 99b = c$, then necessarily $v_2(a), v_2(b), v_2(c)$ are not pairwise distinct (if $v_2(a) \neq v_2(b)$, the strong triangle inequality gives $v_2(c) = \min(v_2(a), v_2(99b)) = \min(v_2(a), v_2(b))$), and the result follows. \square

2. (Iran 2012) Prove that for any positive integer t there is an integer $n > 1$ relatively prime to t such that none of the numbers $n + t, n^2 + t, n^3 + t, \dots$ is a perfect power.

Proof. Let p be a prime divisor of $t + 1$ and let $k = v_p(1 + t)$. We choose n such that $n \equiv 1 \pmod{p^{k+1}}$, then $n^j + t \equiv 1 + t \pmod{p^{k+1}}$ for all

$j \geq 1$, thus $v_p(n^j + t) = k$. Assume that $n^j + t$ is a perfect power, then by the previous relation it must be of the form a^N for some $a > 1$ and $N \mid k$. Take now $n = x^{k!}$ with $x \equiv 1 \pmod{p^{k+1}}$. If $n^j + t = a^N$ then $a^N = t + b^N$, where $b = x^{\frac{k!}{N}}$. Thus $a \geq b + 1$ and then

$$t \geq (b+1)^N - b^N > Nb > x.$$

Choosing $x > t$ and n as above yields therefore a solution of the problem. \square

3. Prove that if n, k are positive integers, then no matter how we choose signs \pm

$$\pm \frac{1}{k} \pm \frac{1}{k+1} \pm \dots \pm \frac{1}{k+n}$$

is not an integer.

Proof. Let

$$r = \max_{k \leq j \leq k+n} v_2(j)$$

and fix $j \in \{k, \dots, k+n\}$ such that $v_2(j) = r$. We claim that there is no other $j' \in \{k, \dots, k+n\}$ such that $v_2(j') = r$. Indeed, otherwise we may assume that $j' > j$, then writing $j = 2^r \cdot a$, $j' = 2^r \cdot b$ with odd integers a, b , we have $k \leq 2^r a < 2^r b \leq k+n$. But then $2^r(a+1)$ belongs to $\{k, \dots, k+n\}$ and $v_2(2^r(a+1)) \geq r+1$, a contradiction.

Thus there is a unique $j \in \{k, \dots, k+n\}$ with $v_2(j) = r$ being maximal. It is now easy to conclude: we can write

$$\pm \frac{1}{k} \pm \frac{1}{k+1} \pm \dots \pm \frac{1}{k+n} = \frac{x}{y} \pm \frac{1}{j}$$

with x, y integers such that $v_2(y) < r$. Since $v_2(j) = r$, it is clear that $\frac{x}{y} \pm \frac{1}{j} = \frac{xj \pm y}{yj}$ cannot be an integer: the numerator is nonzero (as $v_2(y) < v_2(xj)$) and has smaller 2-adic valuation than the denominator! \square

4. (Romania TST 2007) Let $n \geq 3$ and let a_1, \dots, a_n be positive integers such that $\gcd(a_1, \dots, a_n) = 1$ and $\text{lcm}(a_1, \dots, a_n) \mid a_1 + \dots + a_n$. Prove that $a_1 a_2 \dots a_n$ divides $(a_1 + a_2 + \dots + a_n)^{n-2}$.

Proof. It suffices to prove that for any prime p we have

$$v_p(a_1 \dots a_n) \leq (n-2)v_p(a_1 + \dots + a_n).$$

Let $x_i = v_p(a_i)$ and assume without loss of generality that $x_1 \geq \dots \geq x_n$. Since $\gcd(a_1, \dots, a_n) = 1$, we must have $x_n = 0$. Also, by hypothesis

$$v_p(a_1 + \dots + a_n) \geq \max_{1 \leq i \leq n} x_i.$$

If $x_{n-1} \neq 0$, then a_1, \dots, a_{n-1} are multiples of p and $p \mid \text{lcm}(a_1, \dots, a_n) \mid a_1 + \dots + a_n$, a contradiction with the fact that p does not divide a_n . Thus $x_{n-1} = 0$. It is then clear that

$$\begin{aligned} v_p(a_1 \dots a_n) &= x_1 + \dots + x_n = x_1 + \dots + x_{n-2} \leq (n-2) \max_{1 \leq i \leq n} x_i \\ &\leq v_p(a_1 + \dots + a_n), \end{aligned}$$

as desired. □

5. (Erdős-Turan) Let p be an odd prime and let S be a set of n positive integers. Prove that one can choose a subset T of S with at least $\lceil \frac{n}{2} \rceil$ elements such that for all distinct elements $a, b \in T$ we have

$$v_p(a+b) = \min(v_p(a), v_p(b)).$$

Proof. Let $a_1 < \dots < a_n$ be the elements of S . Set $k_i = v_p(a_i)$ and let $a_i = p^{k_i} b_i$ with $b_i > 0$ not divisible by p . Let I (respectively J) be the set of those $i \in \{1, 2, \dots, n\}$ for which the remainder of b_i when divided by p is smaller (respectively larger) than $\frac{p}{2}$. Clearly, if $i, j \in I$ or $i, j \in J$ are distinct then $b_i + b_j$ is not a multiple of p . One of the sets I, J has at least $\lceil \frac{n}{2} \rceil$ elements. Without loss of generality, assume that this set is I . Let $T = \{a_i \mid i \in I\}$. If $i \neq j \in T$ and if $k_i \neq k_j$, then $v_p(a_i + a_j) = \min(v_p(a_i), v_p(a_j))$, so assume that $k_i = k_j$. Then

$$v_p(a_i + a_j) = k_i + v_p(b_i + b_j) = k_i = \min(k_i, k_j),$$

as desired. □

6. (Ostrowski) Find all functions $f: \mathbf{Q} \rightarrow [0, \infty)$ such that

- i) $f(x) = 0$ if and only if $x = 0$;
- ii) $f(xy) = f(x) \cdot f(y)$ and $f(x + y) \leq \max(f(x), f(y))$ for all x, y .

Proof. First of all $f(1) = f(1) \cdot f(1)$ and $f(1) > 0$ (by i)), thus $f(1) = 1$. Similarly, we obtain $f(-1)^2 = 1$ and $f(-1) = 1$. In particular $f(-x) = f(x)$ for all x . Since $f(n) \leq \max(f(n-1), f(1)) = \max(f(n-1), 1)$ for all $n \geq 2$, an immediate induction yields $f(n) \leq 1$ for $n \geq 1$, and since $f(-x) = f(x)$, we obtain $f(x) \leq 1$ for $x \in \mathbf{Z}$.

Suppose first that $f(n) = 1$ for all nonzero integers n . Then for any $x \in \mathbf{Q}^*$ we can find $n \in \mathbf{Z}^*$ such that $nx \in \mathbf{Z}^*$, thus $1 = f(nx) = f(n)f(x) = f(x)$ and $f(x) = 1$ for all nonzero x . This function is indeed a solution of the problem.

Suppose now that there is $n \in \mathbf{Z}^*$ such that $f(n) \neq 1$, i.e. $f(n) < 1$. Take the smallest such positive integer n . Then $n > 1$, and if n is composite, say $n = ab$, then $f(a)f(b) < 1$ forces $f(a) < 1$ or $f(b) < 1$, contradicting the minimality of n . Thus $n = p$ is a prime. We claim that $f(n) = f(p)^{v_p(n)}$ for all n . Since both f and $f(p)^{v_p}$ are totally multiplicative, it suffices to check this for primes n . If $n = p$ this is clear, so assume that n is a prime different from p . Then we can find integers a, b such that $1 = an + bp$, thus

$$1 = f(1) = f(an + bp) \leq \max(f(a)f(n), f(b)f(p)).$$

Since $f(b)f(p) \leq f(p) < 1$, it follows that $f(a)f(n) \geq 1$ and since $f(a), f(n) \leq 1$, we must have $f(a) = f(n) = 1$. Thus $f(n) = 1$ for all primes $n \neq p$ and the claim is proved. We deduce that for any $x = \frac{m}{n}$ we have

$$f(x) = \frac{f(m)}{f(n)} = f(p)^{v_p(m) - v_p(n)}.$$

Conversely, if $a \in (0, 1)$ and p is a prime, setting (for nonzero integers m, n)

$$f\left(\frac{m}{n}\right) = a^{v_p(m) - v_p(n)}$$

yields a solution of the problem, using the basic properties of the v_p -map. \square

7. Find all integers $n > 1$ for which

$$n^n \mid (n-1)^{n^{n+1}} + (n+1)^{n^{n-1}}.$$

Proof. If n is a solution of the problem, then $n \mid (n-1)^{n^{n+1}} + (n+1)^{n^{n-1}}$. If n is even, then

$$(n-1)^{n^{n+1}} + (n+1)^{n^{n-1}} \equiv (-1)^{n^{n+1}} + 1 \equiv 2 \pmod{n},$$

thus $n = 2$ and one easily checks that this is not a solution. So all solutions must be odd. Conversely, we will prove that odd numbers are solutions of the problem, by proving that when n is odd we have

$$n^n \mid (n-1)^{n^{n+1}} + 1 \quad \text{and} \quad n^n \mid (n+1)^{n^{n-1}} - 1.$$

Let p be a prime divisor of n . Then p is odd and by the lifting the exponent lemma

$$v_p((n-1)^{n^{n+1}} + 1) = v_p(n) + v_p(n^{n+1}) = (n+2)v_p(n) > nv_p(n) = v_p(n^n)$$

and

$$v_p((n+1)^{n^{n-1}} - 1) = v_p(n) + v_p(n^{n-1}) = v_p(n^n).$$

The result follows. \square

8. (Mathlinks Contest) Let a, b be distinct positive rational numbers such that $a^n - b^n \in \mathbb{Z}$ for infinitely many positive integers n . Show that $a, b \in \mathbb{Z}$.

Proof. By taking a common denominator of a, b , we are reduced to proving the following statement: if a, b, c are positive integers with $a \neq b$, and if $c^n \mid a^n - b^n$ for infinitely many n , then $c \mid a$ and $c \mid b$. By working with each prime factor of c separately, we easily reduce to the case when c is a prime. Thus we need to prove the following statement: if a, b are

distinct positive integers and p^n divides $a^n - b^n$ for infinitely many n , then p divides a and b . To prove this, assume that p does not divide a , so it does not divide b either. But then

$$\begin{aligned} v_p(a^n - b^n) &\leq v_p(a^{(p-1)n} - b^{(p-1)n}) \leq v_p(a^{2(p-1)} - b^{2(p-1)}) + v_p(n) \\ &\leq c_1 + c_2 \log n \end{aligned}$$

for two constants c_1, c_2 depending only on a, b, p . As the last quantity is smaller than n for large enough n , the result follows. \square

9. (Saint Petersburg) Find all positive integers m, n such that $m^n | n^m - 1$.

Proof. Let p be the smallest prime factor of m . Then p divides $n^{p-1} - 1$ (note that p does not divide n since $p \mid n^m - 1$) and also $n^m - 1$, thus p divides $n^{\gcd(m, p-1)} - 1 = n - 1$. Next, suppose that $p > 2$, then lifting the exponent lemma yields

$$nv_p(m) \leq v_p(n^m - 1) = v_p(n - 1) + v_p(m),$$

from which $n - 1 \leq v_p(n - 1)$, that is $n - 1 \geq p^{n-1} \geq 3^{n-1}$, impossible. So $p = 2$ and n is odd. Then using the lifting the exponent lemma again yields

$$nv_2(m) \leq v_2(n^m - 1) = v_2(n^2 - 1) - 1 + v_2(m),$$

thus $(n - 1)v_2(m) \leq v_2(n^2 - 1) - 1$. It is not difficult to see that this implies $n^2 - 2 \geq 2^{n-1}$ and so that $n = 3$ and $v_2(m) = 1$. Next, m^3 divides $3^m - 1$. Suppose $m > 2$ and let q be the smallest prime factor of $m/2$. Then q is odd and q divides $9^{\gcd(q-1, m/2)} - 1 = 8$, a contradiction. Thus $n = 3$ and $m = 2$. \square

10. (Balkan 1993) Let p be a prime and let $m \geq 2$ be an integer. Prove that if the equation

$$\frac{x^p + y^p}{2} = \left(\frac{x + y}{2} \right)^m$$

has a positive integer solution $(x, y) \neq (1, 1)$, then $m = p$.

Proof. Suppose first that $p = 2$ and $m \geq 3$. Then

$$\frac{(x+y)^2}{2} > \frac{x^2+y^2}{2} = \left(\frac{x+y}{2}\right)^m \geq \left(\frac{x+y}{2}\right)^3,$$

which gives $x+y < 4$. Since $(x, y) \neq (1, 1)$, we must have $x = 1, y = 2$ or $x = 2, y = 1$ and in both cases it is easy to check that the given relation cannot be satisfied.

Assume now that $p > 2$. The function $x \mapsto x^p$ being convex on $[0, \infty)$, Jensen's inequality yields

$$\frac{x^p + y^p}{2} \geq \left(\frac{x+y}{2}\right)^p,$$

which combined with the given relation yields $m \geq p > 2$. Letting $x = du, y = dv$, with $\gcd(u, v) = 1$, the given relation can be written as

$$d^{m-p}(u+v)^m = 2^{m-1}(u^p + v^p).$$

If $u+v$ has an odd prime factor q , then lifting the exponent lemma and the previous equality give

$$mv_q(u+v) = v_q(u^p + v^p) = v_q(p) + v_q(u+v) \leq 1 + v_q(u+v),$$

a contradiction with $m > 2$. Thus $u+v$ is a power of 2 and so

$$v_2(u^p + v^p) = v_2(u+v),$$

since p is odd. But then

$$mv_2(u+v) \leq v_2(u+v) + m - 1,$$

yielding $u+v = 2$, then $u = v = 1$, $x = y$ and finally $m = p$. \square

11. (China TST 2004) Let a be a positive integer. Prove that the equation $n! = a^b - a^c$ has a finite number of solutions (n, b, c) in positive integers.

Proof. Let (n, b, c) be a solution. Note that $a > 1$ and let p be an odd prime not dividing a . Using lifting the exponent lemma and Fermat's little theorem we obtain

$$v_p(a^b - a^c) = v_p(a^{b-c} - 1) \leq v_p((a^{p-1})^{b-c} - 1) = v_p(a^{p-1} - 1) + v_p(b - c).$$

We conclude that

$$v_p(a^{p-1} - 1) + v_p(b - c) \geq v_p(a^b - a^c) = v_p(n!) > \frac{n}{p} - 1.$$

Therefore $v_p(b - c) \geq \frac{n}{p} - k$, for some constant k independent of n .

Letting $\epsilon = p^{-k} > 0$, we conclude that $b - c \geq \epsilon p^{n/p}$ and so

$$n^n > n! = a^b - a^c > a^{b-c} \geq a^{\epsilon p^{n/p}}.$$

Taking logarithms, we deduce that n is bounded in terms of a . Since $c, b - c < n!$, the result follows. \square

12. (China TST 2016) Let c, d be integers greater than 1. Define a sequence $(a_n)_{n \geq 1}$ by $a_1 = c$ and $a_{n+1} = a_n^d + c$ for $n \geq 1$. Prove that for any $n \geq 2$ there is a prime number p dividing a_n and not dividing $a_1 a_2 \dots a_{n-1}$.

Proof. Suppose that for some $n > 1$ there is no such prime p . Take any prime factor p of a_n . By assumption there is some $j < n$ such that $p \mid a_j$. Take the smallest such j . We claim that $j \mid n$. Indeed, $a_{j+1} = a_j^d + c \equiv a_1 \pmod{p}$ and an immediate induction shows that $a_{j+u} \equiv a_u \pmod{p}$ for all $u \geq 1$, i.e. the sequence $(a_n)_{n \geq 1}$ is periodic with period j modulo p . Writing the Euclidean division $n = qj + r$ and assuming that $r > 0$, we obtain $a_r \equiv a_{qj+r} = a_n \equiv 0 \pmod{p}$, thus $p \mid a_r$. This contradicts the minimality of j , thus $j \mid n$.

Next, we claim that $v_p(a_n) = v_p(a_j)$. Let $r = v_p(a_j)$. Then

$$a_{j+1} = a_j^d + c \equiv c = a_1 \pmod{p^{rd}}$$

and again an immediate induction shows that $(a_n)_{n \geq 1}$ becomes periodic with period j modulo p^{rd} . In particular, since $j \mid n$, we have $a_n \equiv a_j \pmod{p^{dr}}$. Since $v_p(a_j) = r < dr$, this gives $v_p(a_n) = r$, as claimed.

The above paragraphs show that for each prime $p \mid a_n$ we can find $j_p < n$ such that $v_p(a_n) = v_p(a_{j_p})$. This implies that a_n divides $a_1 a_2 \dots a_{n-1}$, in particular $a_n \leq a_1 a_2 \dots a_{n-1}$. But an immediate induction shows that $a_n > a_1 \dots a_{n-1}$ for $n \geq 2$. Indeed, this is clear for $n = 2$ and if it holds for n , then the inductive hypothesis combined with the recurrence relation yield (using that $d > 1$)

$$a_{n+1} = a_n^d + c > a_n^d > a_n^{d-1} \cdot a_1 \dots a_{n-1} \geq a_1 a_2 \dots a_n.$$

This contradiction shows that our assumption was wrong, so there is at least one prime dividing a_n and not dividing $a_1 a_2 \dots a_{n-1}$. \square

13. (Kvant M 1687) Find the largest possible number of elements of the set $\{2^n - 1 \mid n \in \mathbf{Z}\}$ that are terms of a geometric progression.

Proof. We will prove that a geometric progression cannot contain more than 2 elements of the set $S = \{2^n - 1 \mid n \in \mathbf{Z}\}$ (which will show that the answer of the problem is 2). Suppose the contrary and set (for some pairwise distinct integers a, b, c)

$$2^a - 1 = a_1 q^\alpha, \quad 2^b - 1 = a_1 q^\beta, \quad 2^c - 1 = a_1 q^\gamma,$$

where $\alpha > \beta > \gamma \geq 0$. Take positive integers m and n , $n > m$, such that $n(\beta - \gamma) = m(\alpha - \gamma)$. Then one easily checks that the previous equalities yield

$$(2^b - 1)^n = (2^a - 1)^m (2^c - 1)^{n-m}.$$

Using the identity $2^x(2^{-x} - 1) = 1 - 2^x$ and taking absolute values we get an equality of the form

$$(2^B - 1)^n = (2^A - 1)^m (2^C - 1)^{n-m},$$

where A, B, C are pairwise distinct positive integers. This immediately implies that $\max(A, C) > B$. On the other hand, the previous equality implies that any prime factor of $2^A - 1$ divides $2^B - 1$ and similarly any prime factor of $2^C - 1$ divides $2^B - 1$. We may assume that $A > B$.

Any prime factor of $2^A - 1$ divides $2^B - 1$ and so it divides $2^d - 1$, where $d = \gcd(A, B) < A$. Since $2^d - 1 \mid 2^A - 1$, it follows that $2^d - 1$ and $2^A - 1$ have the same prime divisors. This contradicts the result established in example 6.31. \square

14. (Iran TST 2009) Let a be a positive integer. Prove that there are infinitely many primes dividing at least one of the numbers

$$2^{2^1} + a, 2^{2^2} + a, 2^{2^3} + a, \dots$$

Proof. We argue by contradiction, assuming that there are only finitely many such primes, say p_1, \dots, p_d . Fix an arbitrary positive integer k and let N be an integer (depending on k) such that for all $n \geq N$ we have

$$2^{2^n} + a > (p_1 \dots p_d)^k.$$

In particular, for any $n \geq N$ there is $i_n \in \{1, \dots, d\}$ such that

$$v_{p_{i_n}}(2^{2^n} + a) > k,$$

since all prime factors of $2^{2^n} + a$ are among p_1, \dots, p_d . Among the $d + 1$ numbers $i_N, i_{N+1}, \dots, i_{N+d}$ (which are all between 1 and d) there must be two equal numbers, say $i_n = i_m$ with $N \leq n < m \leq N + d$. Then $p_{i_n}^k \mid 2^{2^n} + a$ and $p_{i_n}^k = p_{i_m}^k \mid 2^{2^m} + a$. Note that if an integer s divides $2^{2^n} + a$ and $2^{2^m} + a$, then

$$-a \equiv 2^{2^n} = 2^{2^n \cdot 2^{m-n}} \equiv (-a)^{2^{m-n}} \pmod{s}$$

in other words s divides $a^{2^{m-n}} + a$. We deduce that $p_{i_n}^k \mid a^{2^{m-n}} + a$, thus (recall that $N \leq n < m \leq N + d$)

$$2^k \leq p_{i_n}^k \leq a^{2^{m-n}} + a < a^{2^d} + a.$$

Since k was arbitrary and a and d are fixed, this is certainly impossible. The result follows. \square

15. (China TST 2016) A point in the coordinate plane is called rational if its coordinates are rational numbers. Given a positive integer n , can we color all rational points using n colors such that
- each point receives one color;
 - any line segment whose endpoints are rational points contains rational points of each of the n colors?

Proof. Give the color 0 (or any number between 0 and $n-1$) to the origin $O = (0, 0)$. Extend v_2 to \mathbf{Q}^* by setting $v_2(x/y) = v_2(x) - v_2(y)$ for any nonzero integers x, y . This is well-defined and satisfies the same basic properties as v_2 on \mathbf{Z} (this follows immediately from the definition). If $P \neq O$ is a rational point with coordinates x, y , give P the color whose number is the remainder of $\min(v_2(x), v_2(y))$ when divided by n .

Consider now a segment I whose endpoints are $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$. Fix $i \in \{0, 1, \dots, n-1\}$ and let us prove the existence of a point on I whose color is i . We may assume that $v_2(x_1 - x_2) \leq v_2(y_1 - y_2)$, which in particular implies that $x_1 \neq x_2$. Pick k large enough, to ensure that $v_2(x_1 - x_2) - k < v_2(x_2)$ and, if $y_1 \neq y_2$, that $v_2(y_1 - y_2) - k < v_2(y_2)$. Let

$$Q_k = \frac{1}{2^k} P_1 + \left(1 - \frac{1}{2^k}\right) P_2 = \left(\frac{x_1 - x_2}{2^k} + x_2, \frac{y_1 - y_2}{2^k} + y_2\right).$$

Note that Q_k belongs to I and is not equal to O for large enough k . Call u_k, v_k the coordinates of Q_k , so

$$u_k = \frac{x_1 - x_2}{2^k} + x_2, \quad v_k = \frac{y_1 - y_2}{2^k} + y_2.$$

Thanks to the choice of k we have

$$v_2(u_k) = v_2\left(\frac{x_1 - x_2}{2^k}\right) = v_2(x_1 - x_2) - k.$$

If $y_1 = y_2$ then $v_2(v_k) = v_2(y_2)$, while if $y_1 \neq y_2$, then a similar computation shows that $v_2(v_k) = v_2(y_1 - y_2) - k$. In all cases, taking into

account the inequality $v_2(x_1 - x_2) \leq v_2(y_1 - y_2)$ we see that for k large enough we have

$$\min(v_2(u_k), v_2(v_k)) = v_2(u_k) = v_2(x_1 - x_2) - k.$$

It suffices therefore to choose k large enough and such that

$$v_2(x_1 - x_2) - k \equiv i \pmod{n},$$

and then Q_k will receive color i . □

16. (China TST 2010) Let $k > 1$ be an integer and let $n = 2^{k+1}$. Prove that for any positive integers $a_1 < a_2 < \dots < a_n$, the number

$$\prod_{1 \leq i < j \leq n} (a_i + a_j)$$

has at least $k + 1$ different prime divisors.

Proof. We will prove a stronger result, with $n = 2^k + 1$ instead of 2^{k+1} . Suppose that $n = 2^k + 1$ and that $N = \prod_{1 \leq i < j \leq n} (a_i + a_j)$ has at most k prime divisors. Note that N is clearly even, so let $2, q_1, \dots, q_r$ be the different prime factors of N , with $r \leq k - 1$.

By problem 5 we can choose at least $2^{k-1} + 1$ integers $b_1, \dots, b_{2^{k-1}+1}$ among the a_i 's such that $v_{q_1}(b_i + b_j) = \min(v_{q_1}(b_i), v_{q_1}(b_j))$ for $i \neq j$. Note that all prime factors of $\prod_{1 \leq i < j \leq 2^{k-1}+1} (b_i + b_j)$ are among $2, q_1, \dots, q_r$.

Using problem 5 again with the numbers $b_1, \dots, b_{2^{k-1}+1}$, we can find at least $2^{k-2} + 1$ of them, say $c_1, \dots, c_{2^{k-2}+1}$ such that $v_{q_2}(c_i + c_j) = \min(v_{q_2}(c_i), v_{q_2}(c_j))$ for $i \neq j$. Of course, we also have $v_{q_1}(c_i + c_j) = \min(v_{q_1}(c_i), v_{q_1}(c_j))$ for $i \neq j$. Continuing this way we obtain at least 3 numbers x_1, x_2, x_3 among the a_i 's such that

$$v_{q_k}(x_i + x_j) = \min(v_{q_k}(x_i), v_{q_k}(x_j))$$

for all $i \neq j$ and $1 \leq k \leq r$. We will prove that this cannot happen.

Let $x_1 + x_2 = 2^A q_1^{\alpha_1} \dots q_r^{\alpha_r}$ and note that x_1, x_2 are multiples of $q_1^{\alpha_1} \dots q_r^{\alpha_r}$. If $v_2(x_1) \neq v_2(x_2)$, then $A = v_2(x_1 + x_2) = \min(v_2(x_1), v_2(x_2))$ and so x_1, x_2 are also multiples of 2^A , hence x_1, x_2 are multiples of $x_1 + x_2$, impossible. Hence $v_2(x_1) = v_2(x_2)$ and by symmetry we have $v_2(x_1) = v_2(x_2) = v_2(x_3)$. Call this common value B and write $x_i = 2^B y_i$ with y_i odd. Then

$$y_1 + y_2 = 2^{A-B} q_1^{\alpha_1} \dots q_r^{\alpha_r}.$$

Since y_1, y_2 are odd, we have $A - B \geq 1$. Assuming that $A - B = 1$, we deduce that $2x_1, 2x_2$ are both multiples of $x_1 + x_2$, hence $2x_1 \geq x_1 + x_2$ and $2x_2 \geq x_1 + x_2$, forcing $x_1 = x_2$, a contradiction. Thus $A - B \geq 2$ and so $4 \mid y_1 + y_2$. Similarly, $4 \mid y_2 + y_3$ and $4 \mid y_3 + y_1$. This is however impossible, since y_1, y_2, y_3 are odd. \square

17. (Komal) Which binomial coefficients are powers of a prime?

Proof. If $\binom{n}{k} = p^t$, then $p^t \leq n$ by theorem 6.44. Thus $\binom{n}{k} \leq n$. Assume that $2 \leq k \leq n - 2$, then

$$\binom{n}{k} = n \cdot \frac{(n-1)(n-2)\dots(n-k+1)}{k!} \geq n \cdot \frac{(k+1)k \dots \cdot 3}{k!} = n \frac{k+1}{2} > n,$$

a contradiction. Thus $k = 1$ or $k = n - 1$ and the equation reduces to $n = p^t$. \square

18. Prove that $\binom{2n}{n} \mid \text{lcm}(1, 2, \dots, 2n)$ for all positive integers n .

Proof. Let p be any prime and let $k = v_p\left(\binom{2n}{n}\right)$. Then $p^k \leq 2n$ by theorem 6.44, hence $p^k \mid \text{lcm}(1, 2, \dots, 2n)$ since

$$v_p(\text{lcm}(1, 2, \dots, 2n)) = \lfloor \log_p(2n) \rfloor$$

by example 6.7. The result follows. \square

19. Prove that for all positive integers n and all integers a we have

$$\frac{1}{n!}(a^n - 1)(a^n - a)\dots(a^n - a^{n-1}) \in \mathbf{Z}.$$

Proof. We may assume that $n > 1$. It suffices to prove that for every prime p we have

$$v_p(n!) \leq v_p(a^n - 1) + v_p(a^n - a) + \dots + v_p(a^n - a^{n-1}).$$

Theorem 6.49 gives $v_p(n!) \leq \left\lfloor \frac{n-1}{p-1} \right\rfloor$, thus it suffices to prove that

$$\sum_{k=0}^{n-1} v_p(a^n - a^k) \geq \left\lfloor \frac{n-1}{p-1} \right\rfloor.$$

If $p \mid a$, then $v_p(a^n - a^k) \geq k$, thus the left-hand side is at least $\frac{n(n-1)}{2} \geq n-1 \geq \left\lfloor \frac{n-1}{p-1} \right\rfloor$. If p does not divide a , Fermat's little theorem shows that p divides $a^{k(p-1)} - 1$ for $1 \leq k \leq \left\lfloor \frac{n-1}{p-1} \right\rfloor$, thus

$$\sum_{k=0}^{n-1} v_p(a^n - a^k) = \sum_{k=1}^n v_p(a^k - 1) \geq \left\lfloor \frac{n-1}{p-1} \right\rfloor. \quad \square$$

20. Prove that if $k < n$ then

$$n \binom{n-1}{k} \mid \text{lcm}(n, n-1, \dots, n-k).$$

Proof. It suffices to prove that for any prime p we have

$$v_p(n(n-1)\dots(n-k)) \leq v_p(k!) + \max_{n-k \leq j \leq n} v_p(j).$$

Let $N = \max_{n-k \leq j \leq n} v_p(j)$ and consider the numbers $n, n-1, \dots, n-k$. For each $j \leq N$ there are at most $1 + \left\lfloor \frac{k}{p^j} \right\rfloor$ multiples of p^j among them. Indeed, if $k = qp^j + r$ with $0 \leq r < p^j$, then there is exactly one multiple

of p^j among $n - sp^j, n - sp^j - 1, \dots, n - (s+1)p^j + 1$ for each $0 \leq s \leq q-1$, and at most one multiple of p^j among $n - qp^j, \dots, n - k$. Thus

$$v_p(n(n-1)\dots(n-k)) \leq \sum_{j=1}^N \left(1 + \left\lfloor \frac{k}{p^j} \right\rfloor\right) = N + \sum_{j=1}^N \left\lfloor \frac{k}{p^j} \right\rfloor \leq N + v_p(k!),$$

as desired. \square

21. (Mathematical Reflections S 206) Find all integers $n > 1$ having a prime factor p such that $v_p(n!) \mid n-1$.

Proof. Write $n = kp$ and observe that by Legendre's formula

$$v_p(n!) = k + \left\lfloor \frac{n}{p^2} \right\rfloor + \dots \geq k.$$

Hence (recall that $s_p(n)$ is the sum of digits of n in base p)

$$p = \frac{n}{k} > \frac{n-1}{k} \geq \frac{n-1}{v_p(n!)} \geq \frac{n-s_p(n)}{v_p(n!)} = p-1,$$

the last equality being a consequence of theorem 6.49. Since $\frac{n-1}{v_p(n!)}$ is an integer belonging to $[p-1, p)$, it can only be $p-1$ and moreover we must have $s_p(n) = 1$, that is n is a power of a prime. Conversely, if $n = p^k$ for some prime k and some $k \geq 1$, then $v_p(n!) = \frac{n-1}{p-1}$ is a divisor of $n-1$. Thus the solutions are all powers of primes. \square

22. (Romania TST 2015) Let k be an integer greater than 1. When n runs through the integers greater than or equal to k , what is the largest number of divisors of $\binom{n}{k}$ that belong to $\{n-k+1, n-k+2, \dots, n\}$?

Proof. Since

$$\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k!} = \frac{n}{k!}(n-1)\dots(n-k+1),$$

we see that $\binom{n}{k}$ can be divisible by $n-1, n-2, \dots, n-k+1$, by choosing n a multiple of $k!$. If we can prove that $\binom{n}{k}$ is never divisible by all numbers $n, n-1, \dots, n-k+1$, then the answer of the problem is $k-1$.

Fix a prime $p \mid k$ and let

$$u = \max(v_p(n), v_p(n-1), \dots, v_p(n-k+1)).$$

Recall that

$$v_p\left(\binom{n}{k}\right) = \sum_{j=1}^{\infty} x_j, \quad \text{where} \quad x_j = \left\lfloor \frac{n}{p^j} \right\rfloor - \left\lfloor \frac{k}{p^j} \right\rfloor - \left\lfloor \frac{n-k}{p^j} \right\rfloor,$$

and each x_j is either 0 or 1. Note that since $p \mid k$, we have

$$x_1 = \left\lfloor \frac{n}{p} \right\rfloor - \frac{k}{p} - \left\lfloor \frac{n-k}{p} \right\rfloor = 0.$$

Also if $j > u$, then by the definition of u none of the numbers $n, n-1, \dots, n-k+1$ is divisible by p^j . Since $\left\lfloor \frac{m}{p^j} \right\rfloor$ is the number of multiples of p^j that are at most m , we conclude that $\left\lfloor \frac{n}{p^j} \right\rfloor = \left\lfloor \frac{n-k}{p^j} \right\rfloor$. Therefore $x_j = 0$ for $j > u$. Hence

$$v_p\left(\binom{n}{k}\right) = \sum_{j=1}^u x_j \leq u - 1.$$

Thus there must be at least one of the numbers $n, n-1, \dots, n-k+1$ which does not divide $\binom{n}{k}$. \square

23. (Mathematical Reflections O 285) Define a sequence $(a_n)_{n \geq 1}$ by $a_1 = 1$ and $a_{n+1} = 2^n(2^{a_n} - 1)$ for $n \geq 1$. Prove that $n! \mid a_n$ for all $n \geq 1$.

Proof. We will prove by induction on n the following stronger statement: for all primes $p \leq n+1$ we have $v_p(a_n) \geq n-p+1$. Note that this implies the desired result, since for any prime $p \leq n$ theorem 6.49 gives

$$v_p(n!) \leq \frac{n-1}{p-1} = \frac{n-p}{p-1} + 1 \leq n-p+1 \leq v_p(a_n),$$

and this implies that $n! \mid a_n$.

It remains to prove the claim. This is clear for $n = 1$, so assume that it holds for n and let us prove it for $n + 1$. Let $p \leq n + 2$ be a prime and let us prove that $v_p(a_{n+1}) \geq n - p + 2$. This is clear for $p = n + 2$, so assume that $p \leq n + 1$. The result is also clear for $p = 2$, so assume that $p > 2$.

Thanks to the inductive hypothesis and the argument at the end of the first paragraph, we know that $n!$ divides a_n , in particular $p-1 \mid a_n$. Using lifting the exponent lemma (theorem 6.22) and Fermat's little theorem, we obtain

$$\begin{aligned} v_p(a_{n+1}) &= v_p(2^{a_n} - 1) = v_p((2^{p-1})^{\frac{a_n}{p-1}} - 1) \geq v_p(2^{p-1} - 1) + v_p\left(\frac{a_n}{p-1}\right) \\ &\geq 1 + v_p(a_n) \geq 1 + n - p + 1 = n - p + 2, \end{aligned}$$

the last inequality being the inductive hypothesis. This proves the inductive step and finishes the proof. \square

24. (China 2015) For which integers k are there infinitely many positive integers n such that $n + k$ does not divide $\binom{2n}{n}$?

Proof. Since the Catalan numbers are integers (see example 2.54), $k = 1$ is not a solution of the problem. We will prove that any integer $k \neq 1$ is a solution, by constructing for such k infinitely many integers $n \geq 1$ for which $n + k$ does not divide $\binom{2n}{n}$. Suppose first that $k \geq 2$ and let p be a prime factor of k . Choose any j such that $p^j > k$ and let $n = p^j - k$. There are at most $j - 1$ carries when adding n to n , since $2n$ has at most j digits in base p , the last one being 0 (since $p \mid k$). Therefore $p^j = n + k$ cannot divide $\binom{2n}{n}$ by Kummer's theorem. Next, suppose that $k \leq 0$ and choose any odd prime $p > 2|k|$. Letting $n = p - k$, we see that $n + k$ does not divide $\binom{2n}{n}$ since there are no carries when n is added to n in base p . The result follows. \square

Remark 8.20. This result, as well as many other interesting ones appears in the article "Divisors of the Middle Binomial Coefficient" by

Carl Pomerance, published in the American Mathematical Monthly, vol. 122, No. 7 (2015), pp 636-644. The author of the paper also proves the following two interesting results (the proofs are however more technical than that of the problem above): for each $k \geq 1$ almost all integers $n > 0$ (in the sense of asymptotic density) satisfy $n + k \mid \binom{2n}{n}$, while for each $k > 0$ the set of $n > k$ with $n - k \mid \binom{2n}{n}$ is infinite, but with upper density $\leq \frac{1}{3}$.

25. (Romania TST 2007) Find all positive integers x, y such that

$$x^{2007} - y^{2007} = x! - y!.$$

Proof. We will prove that the equation has only the trivial solutions (x, x) , with x is a positive integer. Suppose that $x > y$ satisfy

$$x^{2007} - y^{2007} = x! - y!.$$

Clearly $y > 1$, thus y has a prime divisor p . Considering the equation modulo p we obtain $p \mid x$, therefore (a very weak version of) theorem 6.39 gives

$$2007 \leq v_p(x^{2007} - y^{2007}) = v_p\left(y! \left(\frac{x!}{y!} - 1\right)\right) = v_p(y!) < y.$$

Next, choose any prime $q < 2007$ such that $\gcd(2007, q - 1) = 1$. Then $q \mid x! - y!$, since $y > 2007$. We deduce that $q \mid x^{2007} - y^{2007}$ and since $q \mid x^{q-1} - y^{q-1}$ we obtain $q \mid x - y$. Varying q , we easily obtain $x > y + 2007$, which gives

$$x! - y! = y! \left(\frac{x!}{y!} - 1\right) > 2007! \cdot x(x-1)\dots(x-2006) > x^{2007},$$

a contradiction. The result follows. □

26. a) Prove that for all $n \geq 2$ we have

$$v_2 \left(\binom{4n}{2n} - (-1)^n \binom{2n}{n} \right) = s_2(n) + 2 + 3v_2(n),$$

where $s_2(n)$ is the sum of the digits in the base 2 expansion of n .

b) (AMM E 2640) Find the exponent of 2 in the prime factorization of the number

$$\binom{2^{n+1}}{2^n} - \binom{2^n}{2^{n-1}}.$$

Proof. a) Note that

$$\frac{\binom{4n}{2n}}{\binom{2n}{n}} = \frac{(4n)!}{(2n)!} \cdot \left(\frac{n!}{(2n)!} \right)^2$$

and for all k

$$\frac{(2k)!}{k!} = \frac{2 \cdot 4 \cdot \dots \cdot 2k \cdot 1 \cdot 3 \cdot \dots \cdot (2k-1)}{k!} = 2^k \cdot 1 \cdot 3 \cdot \dots \cdot (2k-1).$$

We conclude that

$$\binom{4n}{2n} - (-1)^n \binom{2n}{n} = \binom{2n}{n} \frac{F_n(4n)}{(2n-1)!!},$$

where $(2n-1)!! = 1 \cdot 3 \cdot \dots \cdot (2n-1)$ and

$$F_n(X) = (X-1) \cdot (X-3) \cdot \dots \cdot (X-2n+1) - (-1) \cdot (-3) \cdot \dots \cdot (1-2n).$$

A brutal expansion shows that

$$F_n(X) = -n^2 X + \frac{1}{6} n(1-4n+3n^2) X^2 + \dots,$$

thus

$$F_n(4n) = -4n^3 + \frac{8}{3} n^3(1-4n+3n^2) + \dots$$

The first term has $v_2(4n^3) = 2 + 3v_2(n)$, the next term has $v_2 \geq 3 + 3v_2(n)$, and all other terms are multiples of $(4n)^3$. The strong triangle inequality gives therefore

$$v_2(F_n(4n)) = 2 + 3v_2(n).$$

Combining this relation with $v_2 \left(\binom{2n}{n} \right) = s_2(n)$ yields the desired result.

b) The answer is $3n$ and this follows directly from part a). \square

27. (China TST 2016) Define a function $f : \mathbf{N} \rightarrow \mathbf{Q}^*$ as follows: write a positive integer $n = 2^k m$ with $k \geq 0$ and m odd, and set $f(n) = m^{1-k}$. Prove that for all $n \geq 1$ the number $f(1)f(2)\dots f(n)$ is an integer divisible by any odd positive integer not exceeding n .

Proof. For each prime p we extend v_p to \mathbf{Q}^* by setting

$$v_p(x/y) = v_p(x) - v_p(y)$$

for all nonzero integers x, y . Fix an odd prime $p \leq n$ and let S_j be the set of those $k \in \{1, 2, \dots, n\}$ for which $v_2(k) \geq j$, i.e. the set of multiples of 2^j among $1, 2, \dots, n$. Note that if $k = 2^r m$ with m odd, then

$$v_p(f(k)) = v_p(m^{1-r}) = (1-r)v_p(m) = (1-r)v_p(k) = (1-v_2(k))v_p(k),$$

the last equality using that p is odd. Thus

$$\begin{aligned} v_p(f(1)f(2)\dots f(n)) &= \sum_{k=1}^n (1-v_2(k))v_p(k) = \sum_{j \geq 0} (1-j) \sum_{v_2(k)=j} v_p(k) \\ &= \sum_{j \geq 0} (1-j) \left(\sum_{k \in S_j} v_p(k) - \sum_{k \in S_{j+1}} v_p(k) \right). \end{aligned}$$

Let

$$x_j = \sum_{k \in S_j} v_p(k) = \sum_{l=1}^{\lfloor \frac{n}{2^j} \rfloor} v_p(2^j l) = v_p\left(\left\lfloor \frac{n}{2^j} \right\rfloor!\right).$$

Then the previous equality yields

$$v_p(f(1)f(2)\dots f(n)) = \sum_{j \geq 0} (1-j)(x_j - x_{j+1})$$

$$= x_0 - x_1 - (x_2 - x_3) - 2(x_3 - x_4) - 3(x_4 - x_5) - \dots = x_0 - x_1 - x_2 - \dots$$

Using Legendre's formula and the previous observations (as well as the easily checked identity $\lfloor \frac{x}{n} \rfloor = \left\lfloor \frac{\lfloor x \rfloor}{n} \right\rfloor$ for each real number x), we end up with

$$v_p(f(1)f(2)\dots f(n)) = \sum_{s \geq 1} \left(\left\lfloor \frac{n}{p^s} \right\rfloor - \sum_{j \geq 1} \left\lfloor \frac{n}{2^j p^s} \right\rfloor \right)$$

It suffices to prove that the last quantity is not less than $N := \lfloor \log_p(n) \rfloor$ (which is the maximal value $v_p(a)$ can take when a varies among odd positive integers not exceeding n). Letting

$$y_s = \left\lfloor \frac{n}{p^s} \right\rfloor - \sum_{j \geq 1} \left\lfloor \frac{n}{2^j p^s} \right\rfloor,$$

it suffices to prove that $y_s \geq 0$ for all s , and that $y_s \geq 1$ for $1 \leq s \leq N$. Fix any $s \geq 1$ and let $x = \left\lfloor \frac{n}{p^s} \right\rfloor$. Using again the equality $\lfloor \frac{x}{n} \rfloor = \left\lfloor \frac{\lfloor x \rfloor}{n} \right\rfloor$ for all $x \in \mathbf{R}$ we obtain

$$y_s = x - \sum_{j \geq 1} \left\lfloor \frac{x}{2^j} \right\rfloor.$$

Since $\left\lfloor \frac{x}{2^j} \right\rfloor \leq \frac{x}{2^j}$ for all j , it is clear that $y_s \geq 0$. Moreover, if $s \leq N$ then $x \geq 1$ and so

$$\sum_{j \geq 1} \left\lfloor \frac{x}{2^j} \right\rfloor < \sum_{j \geq 1} \frac{x}{2^j} = x,$$

the inequality being strict because there is at least one j for which 2^j does not divide x . Thus for $s \leq N$ we have $y_s > 0$ and since y_s is an integer, we finally obtain $y_s \geq 1$. We have just proved that

$$v_p(f(1) \dots f(n)) \geq v_p(a)$$

for each odd prime p and each positive odd integer a not exceeding n . The result follows. \square

28. (IMO Shortlist 2014) If x is a real number, we denote by $\|x\|$ the distance between x and the nearest integer. Prove that if a, b are positive integers, then we can find a prime $p > 2$ and a positive integer k such that

$$\left\| \frac{a}{p^k} \right\| + \left\| \frac{b}{p^k} \right\| + \left\| \frac{a+b}{p^k} \right\| = 1.$$

Proof. Let us start by observing that

$$\|x\| = \left| \left\lfloor x + \frac{1}{2} \right\rfloor - x \right|,$$

since $\left\lfloor x + \frac{1}{2} \right\rfloor$ is the nearest integer to x . Thus $\left\lfloor x + \frac{1}{2} \right\rfloor = x \pm \|x\|$ for all x . We deduce that for all a, b, p, k we have

$$\left\lfloor \frac{a+b}{p^k} + \frac{1}{2} \right\rfloor - \left\lfloor \frac{a}{p^k} + \frac{1}{2} \right\rfloor - \left\lfloor \frac{b}{p^k} + \frac{1}{2} \right\rfloor = \pm \left\| \frac{a+b}{p^k} \right\| \pm \left\| \frac{a}{p^k} \right\| \pm \left\| \frac{b}{p^k} \right\|.$$

Suppose that we manage to prove that the left-hand side is ≥ 1 for some prime $p > 2$ and some $k \geq 1$. Each term in the right-hand side has absolute value $< \frac{1}{2}$, thus the only possibility for the previous equality to happen is that the left-hand side is 1 and all signs are + in the right-hand side, i.e. that the conclusion of the problem is satisfied.

Thus we only need to prove the existence of $p > 2$ and $k \geq 1$ such that the number

$$f(p, k) = \left\lfloor \frac{a+b}{p^k} + \frac{1}{2} \right\rfloor - \left\lfloor \frac{a}{p^k} + \frac{1}{2} \right\rfloor - \left\lfloor \frac{b}{p^k} + \frac{1}{2} \right\rfloor$$

is positive. Using the easily checked identity $\left\lfloor x + \frac{1}{2} \right\rfloor = \lfloor 2x \rfloor - \lfloor x \rfloor$, Legendre's formula gives

$$\begin{aligned} \sum_{k \geq 1} f(p, k) &= \sum_{k \geq 1} \left\lfloor \frac{2(a+b)}{p^k} \right\rfloor - \sum_{k \geq 1} \left\lfloor \frac{a+b}{p^k} \right\rfloor \\ &\quad - \sum_{k \geq 1} \left\lfloor \frac{2a}{p^k} \right\rfloor + \sum_{k \geq 1} \left\lfloor \frac{a}{p^k} \right\rfloor - \sum_{k \geq 1} \left\lfloor \frac{2b}{p^k} \right\rfloor + \sum_{k \geq 1} \left\lfloor \frac{b}{p^k} \right\rfloor \\ &= v_p((2a+2b)!) - v_p((a+b)!) - v_p((2a)!) - v_p((2b)!) + v_p(a!) + v_p(b!) \\ &= v_p \left(\frac{(2a+2b)!a!b!}{(2a)!(2b)!(a+b)!} \right). \end{aligned}$$

Since for all $n \geq 1$

$$\frac{(2n)!}{n!} = 2^n \cdot 1 \cdot 3 \cdot \dots \cdot (2n-1),$$

we can rewrite

$$\begin{aligned} A &:= \frac{(2a+2b)!a!b!}{(2a)!(2b)!(a+b)!} = \frac{1 \cdot 3 \cdot \dots \cdot (2a+2b-1)}{1 \cdot 3 \cdot \dots \cdot (2a-1) \cdot 1 \cdot 3 \cdot \dots \cdot (2b-1)} \\ &= \frac{(2a+1)(2a+3)\dots(2a+2b-1)}{1 \cdot 3 \cdot \dots \cdot (2b-1)}. \end{aligned}$$

This shows that $A > 1$ and that A is a rational number with odd numerator and denominator. Thus there must be an odd prime p such that $v_p(A) > 0$. For such p we obtain $\sum_{k \geq 1} f(p, k) > 0$, thus $f(p, k) \geq 1$ for at least one $k \geq 1$, finishing the proof. \square

29. (Erdős-Palfy-Szegedy theorem) Let a, b be positive integers such that the remainder of a when divided by any prime p does not exceed the remainder of b when divided by p . Prove that $a = b$.

Proof. Taking primes larger than a and b , it is clear that $a \leq b$. Assume that $a < b$ from now on.

Note that if q is a prime factor of $b(b-1)\dots(b-a+1)$, then the remainder of b when divided by q is between 0 and $a-1$, thus $q \leq a$ by assumption of the problem (otherwise the remainder of a when divided by q is $a > a-1$). Thus all prime factors of $b(b-1)\dots(b-a+1)$ are between 1 and a . Let P be the set of primes between 1 and a . For each $p \in P$ and $k \geq 1$ let

$$x(k, p) = \left\lfloor \frac{b}{p^k} \right\rfloor - \left\lfloor \frac{a}{p^k} \right\rfloor - \left\lfloor \frac{b-a}{p^k} \right\rfloor.$$

Then $x(k, p)$ is 0 or 1 for all k, p , and the assumption of the problem implies that $x(1, p) = 0$. Indeed, consider the Euclidean divisions $a = qp + r$ and $b = q'p + r'$, then $r' \geq r$ by assumption, so

$$x(1, p) = q' - q - \left\lfloor q' - q + \frac{r' - r}{p} \right\rfloor = 0.$$

Letting k_p be the largest k for which $x(k, p) = 1$, these observations combined with Legendre's formula yield

$$v_p \left(\binom{b}{a} \right) = x(1, p) + x(2, p) + \dots \leq k_p - 1,$$

thus $\binom{b}{a}$ divides $\prod_{p \in P} p^{k_p-1}$ and so

$$\frac{b(b-1)\dots(b-a+1)}{\prod_{p \in P} p^{k_p}} \mid \frac{a!}{\prod_{p \in P} p}.$$

Next, let us observe that since $x(k_p, p) > 0$, we have

$$\left\lfloor \frac{b}{p^{k_p}} \right\rfloor - \left\lfloor \frac{b-a}{p^{k_p}} \right\rfloor > 0,$$

which means that one of the numbers $b-a+1, b-a+2, \dots, b$ is a multiple of p^{k_p} . Since this happens for all $p \in P$, we deduce that

$$\frac{b(b-1)\dots(b-a+1)}{\prod_{p \in P} p^{k_p}} \geq (b-a+1)^{a-|P|}.$$

Indeed, once the divisions are being made, there are still at least $a - |P|$ factors available at the numerator, and each of them is at least $b-a+1$. On the other hand, clearly

$$\frac{a!}{\prod_{p \in P} p} \leq a^{a-|P|}.$$

We conclude that $b-a+1 \leq a$ and so $2a > b$. However, note that if a, b is a solution of the problem with $a < b$, then so is $b-a$ and b (as the remainder of $b-a$ when divided by any p will be the difference between the remainder of b and that of a , under the stated assumptions). Applying the above reasoning, we also obtain $2(b-a) > b$. But then adding the two relations yields the absurd inequality $2b > 2b$. Thus $a = b$. \square

Remark 8.21. This result is the topic of the article " $a \pmod{p} \leq b \pmod{p}$ for all primes p implies $a = b$ " of P. Erdős, P. P. Palfy and M. Szegedy, published in the American Mathematical Monthly, Vol. 94, No. 2 (1987), pp. 169-170.

30. Prove that there exist two consecutive squares such that there are at least 2000 primes between them.

Proof. Let $k = 2000$. Assuming that between any two consecutive squares there are at most k primes, it follows that the total number of primes between 1 and n is at most $k \cdot (1 + \lfloor \sqrt{n} \rfloor)$. Thus we obtain

$$\pi(n) \leq k \cdot (1 + \lfloor \sqrt{n} \rfloor) \leq 2k\sqrt{n}$$

for all n . On the other hand theorem 6.63 gives

$$\pi(n) \geq \frac{\ln 2}{2} \frac{n}{\ln n},$$

hence we obtain

$$\sqrt{n} \leq \frac{4k}{\ln 2} \ln n,$$

which is impossible for n big enough. \square

31. A finite sequence of consecutive positive integers contains at least one prime number. Prove that the sequence contains a number that is relatively prime to all other terms of the sequence.

Proof. Let $x, x+1, \dots, y$ be the terms of the sequence. Let p be the largest prime appearing in the sequence. Then $2p > y$, for otherwise Bertrand's postulate shows the existence of a prime q between p and $2p$, and such a prime would be in the sequence and greater than p , contradicting the maximality of p . Since $2p > y$, it is clear that p is relatively prime to any term of the sequence which is different from p . The result follows. \square

Remark 8.22. This statement may look innocent, but it actually immediately implies Bertrand's postulate (by applying it to the sequence $2, 3, \dots, 2n$, where $n > 1$ is a given integer), so it is actually equivalent to Bertrand's postulate!

32. Prove that $2p_{n+1} \geq p_n + p_{n+2}$ for infinitely many n , where p_n is the n th prime.

Proof. Assume that this is not the case, say $2p_{n+1} < p_n + p_{n+2}$ for all $n \geq N$ and some $N > 1$. Since $-2p_{n+1} + p_n + p_{n+2}$ is even, we actually

have $p_n + p_{n+2} \geq 2p_{n+1} + 2$ for $n \geq N$. Let $x_n = p_{n+1} - p_n$, then the previous inequality can be written $x_{n+1} \geq x_n + 2$ for $n \geq N$, thus $x_n \geq x_N + 2(n - N) > 2(n - N)$ for $n \geq N$. But then

$$\sum_{k=N}^{n-1} (p_{k+1} - p_k) \geq 2(1 + 2 + \dots + (n - 1 - N)) > (n - 1 - N)^2$$

for $n > N$, hence $p_n > (n - 1 - N)^2$ for $n > N$. This contradicts the result established in example 6.64. \square

33. (AMM) Find all integers $m, n > 1$ such that

$$1! \cdot 3! \cdot \dots \cdot (2n - 1)! = m!.$$

Proof. Clearly $(n, m) = (1, 1), (2, 3)$ are solutions. The equalities

$$3! \cdot 5! = 6!, \quad 3! \cdot 5! \cdot 7! = 720 \cdot 7! = 8 \cdot 9 \cdot 10 \cdot 7! = 10!$$

show that $(n, m) = (3, 6)$ and $(n, m) = (4, 10)$ are also solutions. We will prove that there are no other solutions. Note that if $3! \cdot 5! \cdot 7! \cdot 9! = m!$, then $m! > 11!$ and so $11 \mid m! = 3! \cdot 5! \cdot 7! \cdot 9!$, which is impossible. One proves similarly that $n = 6, 7, 8$ are not solutions. Suppose now that $n \geq 9$. If p is a prime factor of $m!$, then p divides $1! \cdot 3! \cdot \dots \cdot (2n - 1)!$ and so $p \leq 2n - 1$. Thus any prime not exceeding m also does not exceed $2n - 1$. By Bertrand's postulate there is a prime $p \in (\frac{m}{2}, m)$ and so $m < 2(2n - 1)$. On the other hand, we have

$$m > v_2(m!) = \sum_{i=1}^n v_2((2i - 1)!) \geq \sum_{i=1}^n (i - 1) = \frac{n(n - 1)}{2}.$$

We conclude that $n(n - 1) < 4(2n - 1)$, which contradicts the fact that $n > 8$. Thus the only solutions of the problem are

$$(n, m) = (1, 1), (2, 3), (3, 6), (4, 10).$$

\square

34. (EMMO 2016) Let $a_1 < a_2 < \dots$ be an infinite increasing sequence of positive integers such that the sequence $(\frac{a_n}{n})$ is bounded. Prove that for infinitely many n the number a_n divides $\text{lcm}(a_1, \dots, a_{n-1})$.

Proof. Assume by contradiction that a_n does not divide $\text{lcm}(a_1, \dots, a_{n-1})$ for $n \geq N$, for some $N \geq 2$. Thus for each $n \geq N$ we can find a prime p_n dividing a_n such that

$$v_{p_n}(a_n) > v_{p_n}(\text{lcm}(a_1, \dots, a_{n-1})) = \max_{j < n} v_{p_n}(a_j).$$

Let $x_n = p_n^{v_{p_n}(a_n)}$ be the largest power of p_n dividing a_n . Choose $k \geq 1$ such that $a_n \leq kn$ for all $n \geq 1$, thus $x_n \leq kn$ for all n . On the other hand, the inequality above immediately implies that $x_n \neq x_m$ for different $n, m \geq N$. Indeed, if $x_n = x_m$ for some $N \leq n < m$ then $p_n = p_m$ and $v_{p_n}(a_n) = v_{p_m}(a_m)$, contradicting the inequality

$$v_{p_m}(a_m) > v_{p_m}(a_n) = v_{p_n}(a_n).$$

Thus for $n \geq N$ the set $\{1, 2, \dots, kn\}$ contains at least $n - N + 1$ pairwise distinct prime powers, namely x_N, \dots, x_n . On the other hand, it is clear that the number of prime powers in $\{1, 2, \dots, kn\}$ is bounded by

$$\pi(kn) + \sqrt{kn} + \sqrt[3]{kn} + \dots \leq \pi(kn) + \log_2(kn) \cdot \sqrt{kn} \leq c \frac{n}{\log n}$$

for some constant c depending on k (this uses theorem 6.62). We deduce that for $n \geq N$ we have

$$n - N + 1 \leq c \frac{n}{\log n},$$

which is obviously absurd. □

Remark 8.23. As the proof shows, it is enough to ensure that the a_i 's are pairwise distinct and that a_n has order of growth smaller than $n \log n$.

35. Does the equation $x! = y!(y+1)!$ have infinitely many solutions in positive integers?

Proof. The answer is negative. If $x > 8$, choose a prime $q \in (\frac{x}{2}, x]$, which is possible by Bertrand's postulate. Then $v_q(x!) = 1 = 2v_q(y!) + v_q(y+1)$. It follows that $v_q(y!) = 0$ and $v_q(y+1) = 1$, which can only happen when $y = q - 1$. In particular there is a unique such q , namely $y + 1$. Letting $n = \lfloor \frac{x}{2} \rfloor$ we obtain

$$P_n := \prod_{n < p \leq 2n-1} p = q = y + 1 \leq x < 2n + 2.$$

Theorem 6.69 immediately implies that this is impossible for n large enough. Thus x must be bounded and the equation has only finitely many solutions. \square

Remark 8.24. Using the previous argument as well as explicit estimates, it is not difficult (though tedious) to prove that the only solutions in positive integers are given by $(x, y) = (2, 1)$ and $(10, 6)$.

36. (Richert's theorem) Prove that any integer larger than 6 is a sum of distinct primes.

Proof. We will prove by induction on $n \geq 5$ the following statement: each of the numbers $7, 8, \dots, 19 + p_6 + \dots + p_n$ is a sum of distinct primes among p_1, \dots, p_n (here, as usual, p_n denotes the n th prime). For $n = 5$ we need to show that each of the numbers $7, 8, \dots, 19$ is a sum of distinct primes among $2, 3, 5, 7, 11$. This is clear for $7, 8 = 3 + 5, 9 = 2 + 7, 10 = 3 + 7, 11, 12 = 5 + 7, 13 = 2 + 11, 14 = 3 + 11, 16 = 5 + 11, 18 = 7 + 11$ and a little bit less for $15, 17, 19$, which can be written as $3 + 5 + 7, 2 + 3 + 5 + 7, 3 + 5 + 11$ respectively.

Assume now that the statement holds for n and let us prove it for $n + 1$. Write

$$x_n = 19 + p_6 + \dots + p_n.$$

Consider a number $N \in [7, x_{n+1}]$. If $N \in [7, x_n]$, we are done by the inductive hypothesis, so assume that $x_n < N \leq x_{n+1}$. We deduce that

$$x_n - p_{n+1} < N - p_{n+1} \leq x_n.$$

If we manage to prove that $x_n - p_{n+1} \geq 6$, the inductive hypothesis will show that $N - p_{n+1}$ is a sum of distinct primes among p_1, \dots, p_n and so N is a sum of distinct primes among p_1, \dots, p_{n+1} , as needed.

It remains to prove that $x_n - p_{n+1} \geq 6$, which we do by induction. Indeed, for $n = 5$ this is reduced to the equality $19 - 13 = 6$, while if $x_n - p_{n+1} \geq 6$, then thanks to Bertrand's postulate

$$x_{n+1} - p_{n+2} > x_{n+1} - 2p_{n+1} = x_n - p_{n+1} \geq 6. \quad \square$$

Remark 8.25. Schnirelman proved in 1930 that there exists k such that any $n > 1$ is a sum of at most k primes. Riesel and Vaughan proved that every even positive integer is the sum of at most 18 primes, so every integer $n > 1$ is the sum of at most 19 primes.

37. (China TST 2015) Prove that there are infinitely many integers n such that $n^2 + 1$ is squarefree.

Proof. We fix a large integer N and count the numbers n in $\{1, 2, \dots, N\}$ for which $n^2 + 1$ is not squarefree. If n is such a number, then since 4 does not divide $n^2 + 1$, there must be an odd prime p such that $p^2 \mid n^2 + 1$. Then $p^2 \leq n^2 + 1$, so $p \leq N$. Let us fix an odd prime $p \leq N$ and see how many integers $n \in \{1, 2, \dots, N\}$ satisfy $p^2 \mid n^2 + 1$. If n, m are two such integers, then p does not divide mn and $p^2 \mid (m - n)(m + n)$. Since p cannot divide simultaneously $m - n$ and $m + n$, we deduce that $p^2 \mid m - n$ or $p^2 \mid m + n$. Thus any two such integers m, n are either congruent modulo p^2 or their sum is a multiple of p^2 . It is not difficult to deduce that there are at most $2 \left(1 + \frac{N}{p^2}\right)$ such integers n . Therefore the number of $n \in \{1, 2, \dots, N\}$ for which $n^2 + 1$ is squarefree is greater than or equal to

$$N - \sum_{2 < p \leq N} 2 \left(1 + \frac{N}{p^2}\right) > N - 2\pi(N) - 2N \sum_{2 < p \leq N} \frac{1}{p^2}.$$

On the other hand

$$2 \sum_{2 < p \leq N} \frac{1}{p^2} \leq 2 \sum_{k=3}^N \frac{1}{k^2} < 2 \left(\frac{1}{3^2} + \frac{1}{3 \cdot 4} + \frac{1}{4 \cdot 5} + \dots \right) = 2 \left(\frac{1}{3^2} + \frac{1}{3} \right) = \frac{8}{9}.$$

Thus the number of $n \in \{1, 2, \dots, N\}$ for which $n^2 + 1$ is squarefree is greater than

$$N - \frac{8}{9}N - 2\pi(N) = \frac{N}{9} - 2\pi(N).$$

Since $\pi(N) < \frac{N}{10}$ for N large enough, the result follows. \square

38. (USAMO 2014) Prove that there is a constant $c > 0$ with the following property: if a, b, n are positive integers such that $\gcd(a + i, b + j) > 1$ for all $i, j \in \{0, 1, \dots, n\}$, then

$$\min\{a, b\} > c^n \cdot n^{\frac{n}{2}}.$$

Proof. We will prove a stronger statement, with $c^n n^{\frac{n}{2}}$ replaced with $c^n n^n$ for some constant $c > 0$. Note that it is enough to establish such an inequality for n large enough, as then we can always replace c by a smaller constant so that the new inequality holds for all n . Thus n will always be assumed to be sufficiently large.

The idea is fairly simple: consider an $(n + 1) \times (n + 1)$ table and put an arbitrarily chosen prime factor p of $\gcd(a + i, b + j)$ in cell (i, j) for $0 \leq i, j \leq n$. We will prove that at least half of the primes in this table exceed $0.001n^2$ when n is large enough. This will be the technical part of the proof, so take this for granted for a moment and let us see how to conclude. It follows that there is an index $i \in \{0, 1, \dots, n\}$ such that for at least half of the numbers $j \in \{0, 1, \dots, n\}$ the prime in cell (i, j) exceeds $0.001n^2$. All these primes divide $a + i$ and they are pairwise distinct, since if one of these primes divides both $b + j_1$ and $b + j_2$ with $0 \leq j_1 < j_2 \leq n$, then $0.001n^2 < p \leq j_2 - j_1 \leq n$ and this is impossible for n large enough. It follows that $a + i$ is divisible by at least $\frac{n}{2}$ primes, each greater than $0.001n^2$ and so

$$a + n > (0.001n^2)^{\frac{n}{2}},$$

which immediately yields the desired result by symmetry in a and b .

Now, let us prove that at least half of the primes in the table exceed $0.001n^2$. Let $M = \lfloor 0.001n^2 \rfloor$ and let S be the set of primes not exceeding

M. Let us see how many cells can be occupied by a prime $p \in S$. If p occupies cells (i_1, j_1) and (i_2, j_2) , then it divides $i_2 - i_1$ and $j_2 - j_1$ since it divides $a + i_1, a + i_2, b + j_1, b + j_2$. There are at most $1 + \frac{n+1}{p}$ pairwise congruent mod p numbers between 0 and n , hence the number of cells occupied by p does not exceed $\left(1 + \frac{n+1}{p}\right)^2$. It is thus enough to prove that for n large enough we have

$$\sum_{p \in S} \left(1 + \frac{n+1}{p}\right)^2 < \frac{(n+1)^2}{2}.$$

Expand the left-hand side and rewrite the inequality as

$$(n+1)^2 \sum_{p \in S} \frac{1}{p^2} + 2(n+1) \sum_{p \in S} \frac{1}{p} + |S| < \frac{(n+1)^2}{2}.$$

Now, by problem 4.77 the first term in the sum does not exceed $0.49(n+1)^2$. The last term does not exceed $0.001n^2$ by the definition of S . Finally, the second term is bounded by

$$(n+1) \sum_{k=1}^M \frac{1}{k} < (n+1) \log M < 2(n+1) \log n$$

and this is less than $(0.5 - 0.49 - 0.001)n^2$ for n large enough. \square

39. (Mertens) Prove that for all $n > 1$

$$-6 < \sum_{p \leq n} \frac{\ln p}{p} - \ln n < 4.$$

Proof. We will use the prime factorization of $n!$, which yields

$$\log n! = \sum_{p \leq n} v_p(n!) \log p.$$

By theorem 6.39 we have

$$\frac{n}{p} - 1 < v_p(n!) \leq \frac{n}{p-1}.$$

Combining the inequality $v_p(n!) > \frac{n}{p} - 1$ with the obvious inequality $n \log n > \log n!$, we obtain

$$n \log n > \log n! = \sum_{p \leq n} v_p(n!) \log p > n \sum_{p \leq n} \frac{\log p}{p} - \log \prod_{p \leq n} p.$$

Employing Erdős' inequality (theorem 6.57) yields

$$\sum_{p \leq n} \frac{\log p}{p} - \log n < 4,$$

as desired.

Next, similar arguments (using the inequality $v_p(n!) < \frac{n}{p-1}$) yield

$$\frac{\log n!}{n} < \sum_{p \leq n} \frac{\log p}{p} + \sum_{p \leq n} \frac{\log p}{p(p-1)}.$$

In order to bound from above the second sum appearing in the right-hand side, one uses the inequality $\log p < \sqrt{2p}$ (a consequence of the inequality $e^x > \frac{x^2}{2}$ for $x \geq 0$), which gives

$$\sum_{p \leq n} \frac{\log p}{p(p-1)} < \sum_{p \leq n} \frac{\sqrt{2p}}{p(p-1)} \leq \sqrt{2} \sum_{k=2}^n \frac{1}{\sqrt{k}(k-1)}.$$

Finally, we leave to the reader as an amusing exercise to prove the inequality

$$\sum_{k=2}^n \frac{1}{\sqrt{k}(k-1)} < 3.$$

Combining all this yields the desired inequality. □

40. (Mertens) Prove that the sequence $(a_n)_{n \geq 2}$ defined by

$$a_n = \sum_{p \leq n} \frac{1}{p} - \ln \ln n$$

is bounded, where the sum is over all primes not exceeding n .

Proof. The fact that the sequence is bounded from below is an immediate consequence of Euler's theorem 4.74. In order to prove that the sequence is bounded from above we define (for $2 \leq k \leq n$) $u_k = \frac{\log k}{k}$ if k is a prime and $u_k = 0$ otherwise. Letting $S_1 = 0$ and $S_k = u_2 + \dots + u_k$ for $2 \leq k \leq n$, we have

$$\sum_{p \leq n} \frac{1}{p} = \sum_{k=2}^n \frac{u_k}{\ln k} = \sum_{k=2}^n \frac{S_k - S_{k-1}}{\ln k} = \sum_{k=2}^{n-1} S_k \left(\frac{1}{\ln k} - \frac{1}{\ln(k+1)} \right) + \frac{S_n}{\ln n}.$$

By the previous problem

$$S_k = \sum_{p \leq k} \frac{\ln p}{p} < \ln k + 4$$

for $2 \leq k \leq n$. Therefore

$$\sum_{p \leq n} \frac{1}{p} < \frac{4}{\ln 2} + \sum_{k=2}^{n-1} \left(1 - \frac{\ln k}{\ln(k+1)} \right) + \frac{\ln n + 4}{\ln n}.$$

On the other hand for each $4 \leq k \leq n-1$ we have (the last inequality can be proved using the function $f(x) = \ln \ln x$)

$$1 - \frac{\ln k}{\ln(k+1)} = \frac{\ln \left(1 + \frac{1}{k} \right)}{\ln(k+1)} < \frac{1}{k \ln(k+1)} < \frac{1}{k \ln k} < \ln \ln k - \ln \ln(k-1).$$

The result follows now easily by adding the previous inequalities (note that the resulting sum is telescopic). \square

8.6 Congruences for composite moduli

1. (Poland 2003) A polynomial f with integer coefficients has the property that $\gcd(f(a), f(b)) = 1$ for some integers $a \neq b$. Prove that there is an infinite set of integers S such that $\gcd(f(m), f(n)) = 1$ whenever m, n are distinct elements of S .

Proof. It suffices to prove that if a_1, \dots, a_k are integers such that

$$\gcd(f(a_i), f(a_j)) = 1$$

for $1 \leq i \neq j \leq k$ then there is an integer a_{k+1} different from a_1, \dots, a_k and such that $\gcd(f(a_i), f(a_j)) = 1$ for $1 \leq i \neq j \leq k+1$. Pick a_{k+1} such that $a_{k+1} \equiv a_i \pmod{f(a_{i+1})}$ for $1 \leq i \leq k-1$ and $a_{k+1} \equiv a_k \pmod{f(a_1)}$, which is possible by the Chinese remainder theorem. Then $f(a_{k+1}) \equiv f(a_i) \pmod{f(a_{i+1})}$ for $1 \leq i < k$ and $f(a_{k+1}) \equiv f(a_k) \pmod{f(a_1)}$. Thus $\gcd(f(a_{k+1}), f(a_{i+1})) = \gcd(f(a_i), f(a_{i+1})) = 1$ for $1 \leq i < k$ and similarly $\gcd(f(a_{k+1}), f(a_1)) = 1$, as desired. \square

Remark 8.26. In particular, consider two relatively prime integers a, b . Then the problem implies that in the arithmetic progression $(an + b)_{n \geq 0}$ there is an infinite set of pairwise relatively prime numbers. This is of course a direct consequence of Dirichlet's theorem on primes in arithmetic progressions, but as the problem shows one can also prove it by purely elementary means (and with an argument which generalizes to higher degrees, for which no analogue of Dirichlet's theorem is known).

2. Prove that for all positive integers k and n there exists a set S of n consecutive positive integers such that each $x \in S$ has at least k distinct prime divisors that do not divide any other element of S .

Proof. Consider a matrix $(p_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq k}}$ with n rows and k columns and whose entries are pairwise distinct primes greater than n . Let R_1, \dots, R_n be the products of the entries in rows $1, 2, \dots, n$ of the matrix. Then R_1, \dots, R_n are pairwise relatively prime, so by the Chinese remainder theorem there is a positive integer x such that $x \equiv -i \pmod{R_i}$ for $1 \leq i \leq n$. Then $x + i$ has at least k distinct prime divisors (namely the entries of the k th row of the matrix) and none of these divides $x + j$ for $j \neq i$: if p is an entry of the matrix and $p \mid x + i$ and $p \mid x + j$ then $p \mid j - i$, contradicting the fact that $p > n$. Thus $x + 1, \dots, x + n$ satisfy all required properties. \square

3. A lattice point is called visible if its coordinates are relatively prime integers. Prove that for any positive integer k there is a lattice point whose distance from each visible lattice point is greater than k .

Proof. It suffices to prove that for each k we can find a square of side length k with sides parallel to the coordinate axes and consisting of invisible points. In other words, we want to find distinct integers x, y such that $\gcd(x+i, y+j) > 1$ for $1 \leq i, j \leq k$. Consider a $k \times k$ matrix whose entries $(p_{ij})_{1 \leq i, j \leq k}$ are pairwise distinct primes, and let R_1, \dots, R_k (respectively C_1, \dots, C_k) be the products of the numbers in rows (respectively columns) $1, 2, \dots, k$ of the matrix. Then R_1, \dots, R_k (respectively C_1, \dots, C_k) are pairwise relatively prime, so by the Chinese remainder theorem there are integers $x \neq y$ such that $x \equiv -i \pmod{R_i}$ for $1 \leq i \leq k$ and $y \equiv -j \pmod{C_j}$ for $1 \leq j \leq k$. Then for all $1 \leq i, j \leq k$ the prime p_{ij} divides both $x+i$ and $y+j$, hence $\gcd(x+i, y+j) > 1$ and we are done. \square

4. a) Prove that for all $n \geq 1$ there is a positive integer a such that $a, 2a, \dots, na$ are all perfect powers.
- b) (Balkan 2000) Prove that for all $n \geq 1$ there is a set A of n positive integers such that for all $1 \leq k \leq n$ and all $x_1, x_2, \dots, x_k \in A$ the number $\frac{x_1 + x_2 + \dots + x_k}{k}$ is a perfect power.

Proof. a) Choose pairwise distinct primes p_1, \dots, p_n . We will prove that there is $a > 1$ such that ia is a p_i th power for $1 \leq i \leq n$. Letting q_1, \dots, q_k be all primes not exceeding n , we can write each $1 \leq i \leq n$ as $i = q_1^{\alpha_{i1}} \dots q_k^{\alpha_{ik}}$ with $\alpha_{ij} \geq 0$. We look for a of the form $a = q_1^{x_1} \dots q_k^{x_k}$. If $1 \leq i \leq n$, then $ia = q_1^{\alpha_{i1} + x_1} \dots q_k^{\alpha_{ik} + x_k}$ is a p_i th power if $\alpha_{ij} + x_j \equiv 0 \pmod{p_i}$ for $1 \leq j \leq k$. By the Chinese remainder theorem we can choose for each $1 \leq j \leq k$ a positive integer x_j such that $x_j \equiv -\alpha_{ij} \pmod{p_i}$ for $1 \leq i \leq k$ and the problem is solved.

Let us remark that there is also a very simple inductive proof: we prove the result by induction on n , taking $a = 4$ for $n = 1$. Suppose that we can find a such that $ka = x_k^{y_k}$ for $1 \leq k \leq n$, where x_k, y_k are

integers greater than 1. Let m be a common multiple of y_1, \dots, y_n and choose $b = (n+1)^m a^{m+1}$. For $1 \leq k \leq n$ the number kb is a y_k th power since $y_k \mid m$ and since ka is an y_k th power. On the other hand $(n+1)b = ((n+1)a)^{m+1}$ is also a perfect power, so we are done.

b) By part a) there is a positive integer a such that $a, 2a, \dots, n \cdot n!a$ are all perfect powers. Consider the set $A = \{n!a, 2n!a, \dots, nn!a\}$. If $x_1, \dots, x_k \in A$ and $1 \leq k \leq n$, then $\frac{x_1 + \dots + x_k}{k}$ is of the form $\frac{n!}{k} am$ with $1 \leq m \leq nk$. Thus $\frac{x_1 + \dots + x_k}{k}$ is indeed a perfect power by the choice of a . \square

5. Let a, b, c be pairwise distinct positive integers. Prove that there is an integer n such that $a+n, b+n, c+n$ are pairwise relatively prime.

Proof. It is not difficult to see that there is k such that at least two of the numbers $a+k, b+k, c+k$ are odd. Replacing a, b, c with $a+k, b+k, c+k$ and making a permutation of these numbers, we may assume that a and b are odd. Let p_1, \dots, p_m be the odd prime divisors of $(a-b)(b-c)(c-a)$ (we allow $m = 0$). For all $1 \leq i \leq m$ the numbers a, b, c give at most two different remainders when divided by p_i (since p_i divides $a-b$ or $b-c$ or $c-a$), thus (since $p_i > 2$) there is an integer n_i such that $a+n_i, b+n_i, c+n_i$ are not multiples of p_i . Using the Chinese remainder theorem, we can find an even integer n such that $n \equiv n_i \pmod{p_i}$ for $1 \leq i \leq m$. Then $n+a, n+b, n+c$ are pairwise relatively prime: by construction the only possible common prime factor of any two of the numbers $n+a, n+b, n+c$ is 2 (note that any such prime factor would divide $(a-b)(b-c)(c-a)$), which is excluded since $n+a$ and $n+b$ are odd. The result follows. \square

6. (AMM) Prove that there are arbitrarily long sequences of consecutive integers, none of which can be written as the sum of two perfect squares.

Proof. Letting n be a positive integer, we will construct a positive integer x such that none of the numbers $x+1, \dots, x+n$ is a sum of two squares. For this, we choose pairwise distinct primes p_1, \dots, p_n that are congruent

to 3 modulo 4 (this is possible thanks to example 4.56), and then use the Chinese remainder theorem to find x such that

$$x + 1 \equiv p_1 \pmod{p_1^2}, x + 2 \equiv p_2 \pmod{p_2^2}, \dots, x + n \equiv p_n \pmod{p_n^2}.$$

By theorem 5.60 none of the numbers $x + 1, \dots, x + n$ is a sum of two squares. \square

7. Let f be a nonconstant polynomial with integer coefficients and let n and k be positive integers. Prove that there is a positive integer a such that each of the numbers $f(a), f(a+1), \dots, f(a+n-1)$ has at least k distinct prime divisors.

Proof. Choose pairwise distinct prime numbers $(p_{ij})_{1 \leq i, j \leq k}$ such that $f(x_{ij}) \equiv 0 \pmod{p_{ij}}$ for some positive integers x_{ij} , which is possible using Schur's theorem 4.67. Thanks to the Chinese remainder theorem, we can find $a \geq 1$ such that $a+i-1 \equiv x_{ij} \pmod{p_{ij}}$ for all i, j . But then each of the numbers $f(a), f(a+1), \dots, f(a+n-1)$ has at least k distinct prime divisors, since p_{ij} divides $f(a+i-1)$ for all $1 \leq i, j \leq k$. \square

8. (IMC 2013) Let p and q be relatively prime positive integers. Prove that

$$\sum_{k=0}^{pq-1} (-1)^{\left\lfloor \frac{k}{p} \right\rfloor + \left\lfloor \frac{k}{q} \right\rfloor} = \begin{cases} 0 & \text{if } pq \text{ is even} \\ 1 & \text{if } pq \text{ odd} \end{cases}$$

Proof. Write

$$f(k) = \left\lfloor \frac{k}{p} \right\rfloor + \left\lfloor \frac{k}{q} \right\rfloor$$

for $0 \leq k \leq pq-1$. Suppose first that pq is even. If $k \in \{0, 1, \dots, pq-1\}$, then writing $k = \alpha p + r$ with $0 \leq r < p$ we obtain

$$\left\lfloor \frac{pq-1-k}{p} \right\rfloor = q + \left\lfloor \frac{-k-1}{p} \right\rfloor = q - \alpha - 1 = q - 1 - \left\lfloor \frac{k}{p} \right\rfloor$$

and a similar formula with p replaced by q . Since $p + q$ must be odd in this case, it follows that

$$f(pq - 1 - k) = q + p - 2 - f(k) \equiv 1 + f(k) \pmod{2},$$

which immediately yields

$$\sum_{k=0}^{pq-1} (-1)^{f(k)} = \sum_{k=0}^{pq-1} (-1)^{f(pq-1-k)} = - \sum_{k=0}^{pq-1} (-1)^{f(k)}$$

and then $\sum_{k=0}^{pq-1} (-1)^{f(k)} = 0$.

Suppose now that pq is odd. Writing the Euclidean division of $k \in \{0, 1, \dots, pq - 1\}$ by p and q in the form $k = a_k p + r_k$ and $k = b_k q + R_k$, we obtain (since p and q are odd)

$$f(k) = a_k + b_k \equiv k - r_k + k - R_k \equiv r_k + R_k \pmod{2}.$$

Next, for each $(r, R) \in \{0, 1, \dots, p - 1\} \times \{0, 1, \dots, q - 1\}$ the Chinese remainder theorem yields the existence of a unique $k \in \{0, 1, \dots, pq - 1\}$ such that $(r_k, R_k) = (a, b)$. We conclude that

$$\sum_{k=0}^{pq-1} (-1)^{f(k)} = \sum_{a=0}^{p-1} \sum_{b=0}^{q-1} (-1)^{a+b} = \sum_{a=0}^{p-1} (-1)^a \cdot \sum_{b=0}^{q-1} (-1)^b = 1.$$

The result follows. □

9. (IMO 1999 Shortlist) Find all positive integers n for which there is an integer m such that $2^n - 1 \mid m^2 + 9$.

Proof. Clearly $n = 1$ is a solution of the problem, so assume from now on that $n \geq 2$. If $2^n - 1 \mid m^2 + 9$ for some integer m , then $2^n - 1$ has no prime divisor $p > 3$ such that $p \equiv 3 \pmod{4}$, by corollary 5.28. On the other hand, if $d > 1$ is an odd integer, then $2^d - 1 \equiv -1 \pmod{4}$ and 3 does not divide $2^d - 1$, thus $2^d - 1$ has a prime factor $p \equiv 3 \pmod{4}$ different from 3. We deduce that n cannot have any odd divisor $d > 1$ (as

otherwise $2^d - 1 \mid 2^n - 1 \mid m^2 + 9$, contradicting the previous observations) and so n is a power of 2.

Conversely if n is a power of 2, then n is a solution of the problem. Indeed, write $n = 2^k$ and observe that

$$2^n - 1 = 3 \cdot (2^2 + 1)(2^4 + 1) \dots (2^{2^{k-1}} + 1).$$

Choosing $m = 3a$, it is enough to find a such that

$$(2^2 + 1)(2^4 + 1) \dots (2^{2^{k-1}} + 1) \text{ divides } a^2 + 1.$$

Since the Fermat numbers $2^{2^i} + 1$ are pairwise relatively prime (by example 3.12), by the Chinese remainder theorem there is an integer a such that $a \equiv 2^{2^{i-1}} \pmod{2^{2^i} + 1}$ for $1 \leq i \leq k-1$. Then $a^2 + 1 \equiv 0 \pmod{2^{2^i} + 1}$ for $1 \leq i \leq k-1$ and so $(2^2 + 1)(2^4 + 1) \dots (2^{2^{k-1}} + 1)$ divides $a^2 + 1$. The solutions of the problem are therefore all powers of 2. \square

10. (Bulgaria 2003) A finite set C of positive integers is called good if for any $k \in \mathbf{Z}$ there exist $a \neq b \in C$ such that $\gcd(a+k, b+k) > 1$. Prove that if the sum of the elements of a good set C equals 2003, then there exists $c \in C$ such that the set $C - \{c\}$ is good.

Proof. Say a prime p is good for a set $C \subset \mathbf{Z}$ if for any $i \in \{0, 1, \dots, p-1\}$ the congruence $x \equiv i \pmod{p}$ is satisfied by at least two elements of C . It is clear that if there is a prime p good for C , then C is good. The crucial remark is that the converse holds. Indeed, assume that C is good but no prime p is good for C . Let S be the set of all primes not exceeding $\max(C)$. Then for all $p \in S$ we can find $i_p \in \{0, 1, \dots, p-1\}$ such that $x \equiv i_p \pmod{p}$ for at most one element x of C . Using the Chinese remainder theorem, we can find an integer k such that $k \equiv -i_p \pmod{p}$ for all $p \in S$. Since C is good, there are $a \neq b \in C$ such that $\gcd(a+k, b+k) > 1$. Letting p be a prime divisor of $\gcd(a+k, b+k)$, we have $p \in S$, since $p \mid b-a$ and $|a-b| < \max(C)$. But then $a \equiv -k \equiv i_p \pmod{p}$ and $b \equiv i_p \pmod{p}$, contradicting the choice of i_p . This establishes the claim.

It is now rather easy to solve the problem. Let p be a prime good for C , so we can find $a_i \neq b_i \in C$ such that $a_i \equiv b_i \equiv i \pmod{p}$ for $0 \leq i < p$. Note that $\{a_0, \dots, a_{p-1}, b_0, \dots, b_{p-1}\}$ is a good set and so is any set containing it (as p is good for any such set). Thus it suffices to prove that $C \neq \{a_0, \dots, a_{p-1}, b_0, \dots, b_{p-1}\}$. If $C = \{a_0, \dots, a_{p-1}, b_0, \dots, b_{p-1}\}$, then the hypothesis yields

$$\sum_{i=0}^{p-1} (a_i + b_i) = 2003.$$

On the other hand

$$\sum_{i=0}^{p-1} (a_i + b_i) \equiv 2 \sum_{i=0}^{p-1} i \equiv 0 \pmod{p},$$

thus p divides 2003 and then $p = 2003$. Since $a_i, b_i \geq 1$, this makes the equality $\sum_{i=0}^{p-1} (a_i + b_i) = 2003$ impossible and finishes the proof. \square

11. Is there a sequence of 101 consecutive odd integers such that each term of the sequence has a prime factor not exceeding 103?

Proof. Such a sequence exists. Let d be the product of all odd primes not exceeding 47. Let x be an odd positive multiple of d and consider the sequence of 101 consecutive odd integers

$$x - 100, x - 98, x - 96, \dots, x, x + 2, \dots, x + 100.$$

If $1 \leq j \leq 50$ is not a power of 2, then the terms $x \pm 2j$ of this sequence are not relatively prime to d (since j has an odd prime factor, which divides d). There are still 12 terms of the sequence we have to take care of, namely $x - 2^j$ and $x + 2^j$ for $1 \leq j \leq 6$. Call these terms $x - a_1, \dots, x - a_{12}$ and let $p_1 = 53, \dots, p_{12} = 103$ be all primes in $(47, 103]$. Choosing x such that $x \equiv a_i \pmod{p_i}$ for $1 \leq i \leq 12$, which is possible by the Chinese remainder theorem, yields a sequence which has the desired property. \square

12. (USA TST 2010) The sequence $(a_n)_{n \geq 1}$ satisfies $a_1 = 1$ and

$$a_n = a_{\lfloor n/2 \rfloor} + a_{\lfloor n/3 \rfloor} + \dots + a_{\lfloor n/n \rfloor} + 1$$

for all $n \geq 2$. Prove that $a_n \equiv n \pmod{2^{2010}}$ for infinitely many n .

Proof. For $n \geq 3$ we have

$$a_n - a_{n-1} = \sum_{k=2}^n a_{\lfloor \frac{n}{k} \rfloor} - \sum_{k=2}^{n-1} a_{\lfloor \frac{n-1}{k} \rfloor} = 1 + \sum_{k=2}^{n-1} (a_{\lfloor \frac{n}{k} \rfloor} - a_{\lfloor \frac{n-1}{k} \rfloor}).$$

Since $\lfloor \frac{n}{k} \rfloor - \lfloor \frac{n-1}{k} \rfloor$ is 0 unless $k \mid n$, in which case it equals 1, we deduce that

$$a_n - a_{n-1} = 1 + \sum_{\substack{2 \leq d < n \\ d \mid n}} (a_d - a_{d-1}),$$

which also holds for $n = 2$. Defining $x_1 = 1$ and $x_n = a_n - a_{n-1}$ for $n > 1$, the previous relation becomes (for $n \geq 2$)

$$x_n = \sum_{\substack{d \mid n \\ d < n}} x_d.$$

Next we prove by strong induction on n that

$$2^{\max_{p \mid n} v_p(n)-1} \mid x_n.$$

This is immediate for $n = 1$ and $n = 2$, so assume that $n > 2$ and that the previous divisibility holds for $1, 2, \dots, n-1$. Take any prime $p \mid n$ and let $k = v_p(n)$. We have (assuming $n \neq p$, as otherwise everything is clear)

$$x_n = \sum_{\substack{d \mid \frac{n}{p}}} x_d + \sum_{\substack{p^k \mid d \\ d < n}} x_d = 2x_{\frac{n}{p}} + \sum_{\substack{p^k \mid d \\ d < n}} x_d.$$

By the inductive hypothesis 2^{k-1} divides $2x_{n/p}$ and also x_d for all $d < n$ such that $p^k \mid d$. Thus $2^{k-1} \mid x_n$ and the inductive step is proved.

It is now easy to conclude. Let $N = 2^{2010} - 1$ and choose pairwise distinct primes p_1, \dots, p_N . By the Chinese remainder theorem there are infinitely many integers $z > 1$ such that $z \equiv -i \pmod{p_i^{2011}}$ for all $1 \leq i \leq N$. The previous paragraph shows that $2^{2010} \mid x_{z+i}$ for $1 \leq i \leq N$, thus $a_{z+i} \equiv a_{z+i-1} \pmod{2^{2010}}$ for $1 \leq i \leq N$. Since one of the numbers $z, z+1, \dots, z+N$ is congruent to a_z modulo 2^{2010} , it follows that for each such z we can find $n \in \{z, z+1, \dots, z+2^{2010}-1\}$ such that $a_n \equiv n \pmod{2^{2010}}$, which finishes the proof. \square

13. (China TST 2014) A function $f: \mathbf{N} \rightarrow \mathbf{N}$ satisfies for all $m, n \geq 1$

$$\gcd(f(m), f(n)) \leq \gcd(m, n)^{2014} \quad \text{and} \quad n \leq f(n) \leq n + 2014.$$

Prove that there is a positive integer N such that $f(n) = n$ for $n \geq N$.

Proof. Letting $k = 2014$, we will prove that we can take $N = k^k$. First, let us observe that f is injective on (k^k, ∞) , for if $f(a) = f(b)$ with $a > b > k^k$, then $a \leq f(a) = f(b) \leq b + k$, hence

$$a \leq f(a) = \gcd(f(a), f(b)) \leq \gcd(a, b)^k \leq (a - b)^k = k^k,$$

a contradiction. This observation combined with the inequality $n \leq f(n) \leq n + k$ yields the following result: if $x > k^k$ and $f(x+i) = x+i$ for all $1 \leq i \leq k$, then $f(n) = n$ for all $n \in (k^k, x]$. It is thus sufficient to prove the existence of infinitely many x such that $f(x+i) = x+i$ for all $1 \leq i \leq k$.

Let $n > 1$ be an integer such that $f(n) \neq n$. By example 7.8 we know that all prime divisors p of $f(n)$ divide n and so they divide $f(n) - n$. Since $n + 1 \leq f(n) \leq n + k$, it follows that any such prime p cannot exceed k . Calling a number $n > 1$ nice if it has at least one prime factor greater than k , it follows that $f(n) = n$ whenever $f(n)$ is nice. Thus, in order to finish the proof it suffices to prove that for infinitely many x each of the numbers $x+1, \dots, x+k$ is nice. This is however an immediate consequence of the Chinese remainder theorem: simply choose different primes q_1, \dots, q_k each greater than k , and choose x such that $x+i \equiv 0 \pmod{q_i}$ for $1 \leq i \leq k$. \square

14. (Iran 2007) Let n be a positive integer such that $\gcd(n, 2(2^{1386} - 1)) = 1$. Let $a_1, a_2, \dots, a_{\varphi(n)}$ be a reduced residue system modulo n . Prove that

$$n \mid a_1^{1386} + a_2^{1386} + \dots + a_{\varphi(n)}^{1386}$$

Proof. Since n is odd, $2a_1, 2a_2, \dots, 2a_{\varphi(n)}$ is also a reduced residue system modulo n , hence

$$a_1^{1386} + a_2^{1386} + \dots + a_{\varphi(n)}^{1386} \equiv (2a_1)^{1386} + (2a_2)^{1386} + \dots + (2a_{\varphi(n)})^{1386} \pmod{n}.$$

The result follows, since n is relatively prime to $2^{1386} - 1$. \square

15. Let $n > 1$ be an integer and let $r_1, r_2, \dots, r_{\varphi(n)}$ be a reduced residue system modulo n . For which integers a is $r_1 + a, r_2 + a, \dots, r_{\varphi(n)} + a$ a reduced residue system modulo n ?

Proof. Note that $r_1 + a, \dots, r_{\varphi(n)} + a$ are pairwise distinct modulo n , since $r_1, \dots, r_{\varphi(n)}$ are so. So the question is when they are all relatively prime to n . Let $N = \prod_{p|n} p$ be the product of the different prime factors of n . If $N \mid a$, then clearly $\gcd(x + a, n) = 1$ whenever $\gcd(x, n) = 1$, so any such a is a solution of the problem. Conversely, let a be a solution and assume that N does not divide a , so there is a prime factor p of n which does not divide a . Write $n = p^r k$ with k relatively prime to p and choose, using the Chinese remainder theorem, x such that $x \equiv -a \pmod{p}$ and $x \equiv 1 \pmod{k}$. Then clearly $\gcd(x, n) = 1$, so there is i such that $x \equiv r_i \pmod{n}$. It follows that $x + a \equiv r_i + a \pmod{n}$ and so $\gcd(x + a, n) = \gcd(r_i + a, n) = 1$. This is clearly absurd, since $p \mid \gcd(x + a, n)$. Thus the answer of the problem is: all multiples of $\prod_{p|n} p$. \square

16. Prove that any positive integer n has a multiple whose sum of digits is n .

Proof. Write $n = 2^a \cdot 5^b \cdot m$, with $\gcd(m, 10) = 1$ and $a, b \geq 0$, and consider the number

$$A = 10^{a+b}(10^{\varphi(m)} + 10^{2\varphi(m)} + \dots + 10^{n\varphi(m)}).$$

Clearly the sum of digits of A is n . It remains to check that A is a multiple of n . It suffices to prove that m divides $10^{\varphi(m)} + 10^{2\varphi(m)} + \dots + 10^{n\varphi(m)}$, which follows from Euler's theorem. \square

17. For which integers $n > 1$ is there a polynomial f with integer coefficients such that $f(k) \equiv 0 \pmod{n}$ or $f(k) \equiv 1 \pmod{n}$ for any integer k , and both these congruences have solutions?

Proof. If n is a power of a prime, then Euler's theorem shows that $X^{\varphi(n)}$ is a solution of the problem. If n is not a prime power, we can write $n = ab$ with $a, b > 1$ and $\gcd(a, b) = 1$. Fix some r, s with $f(r) \equiv 0 \pmod{n}$ and $f(s) \equiv 1 \pmod{n}$. The Chinese remainder theorem gives a t with $t \equiv r \pmod{a}$ and $t \equiv s \pmod{b}$. But then we see that $f(t) \equiv f(r) \equiv 0 \pmod{a}$ and $f(t) \equiv f(s) \equiv 1 \pmod{b}$. But then $f(t)$ is neither 0 nor 1 modulo n , a contradiction. Thus the solutions of the problem are exactly the powers of prime numbers. \square

18. (Saint Petersburg 1998) Is there a nonconstant polynomial f with integer coefficients and an integer $a > 1$ such that the numbers $f(a), f(a^2), f(a^3), \dots$ are pairwise relatively prime?

Proof. The answer is negative. Assume that $f(a), f(a^2), f(a^3), \dots$ are pairwise relatively prime. Then $\gcd(a, f(0))$ divides both $f(a)$ and $f(a^2)$, so $\gcd(a, f(0)) = 1$. Choose a positive integer i such that $|f(a^i)| > 2$, which is possible since f is nonconstant. Note that $\gcd(a, f(a^i)) = \gcd(a, f(0)) = 1$, thus, letting $j = i + \varphi(|f(a^i)|)$, Euler's theorem yields $a^j \equiv a^i \pmod{f(a^i)}$ and $f(a^j) \equiv f(a^i) \equiv 0 \pmod{f(a^i)}$. Thus $\gcd(f(a^i), f(a^j)) \neq 1$, a contradiction. \square

19. a) (IMO 1971) Prove that the sequence $(2^n - 3)_{n \geq 1}$ contains an infinite subsequence in which every two distinct terms are relatively prime.
b) (Romania TST 1997) Let $a > 1$ be a positive integer. Prove the same result as in a) for the sequence $(a^{n+1} + a^n - 1)_{n \geq 1}$.

Proof. a) We prove by induction on $k \geq 2$ the existence of an increasing sequence $n_1 < \dots < n_k$ such that $\gcd(2^{n_i} - 3, 2^{n_j} - 3) = 1$ for all $i \neq j \in \{1, 2, \dots, k\}$. For $k = 2$ take $n_1 = 1$ and $n_2 = 2$. Assuming that $n_1 < \dots < n_k$ are constructed, we will construct $n_{k+1} > n_k$ such that $2^{n_{k+1}} - 3$ is relatively prime to $N := \prod_{j=1}^k (2^{n_j} - 3)$. Simply take $n_{k+1} = (n_k + 1)\varphi(N)$ and observe that $n_{k+1} > n_k$ and by Euler's theorem $2^{n_{k+1}} - 3 \equiv -2 \pmod{N}$. Since N is odd, it follows that $2^{n_{k+1}} - 3$ is relatively prime to N , as desired.

b) Let $x_n = a^{n+1} + a^n - 1$. As in part a), we prove by induction on $k \geq 1$ the existence of a sequence $n_1 < \dots < n_k$ such that x_{n_1}, \dots, x_{n_k} are pairwise relatively prime. There is nothing to be done for $k = 1$, so assume the existence of $n_1 < \dots < n_k$. Let $N = x_{n_1} \dots x_{n_k}$ and note that $\gcd(N, a) = 1$. Choose $n_{k+1} = (n_k + 1)\varphi(N)$, then by Euler's theorem $x_{n_{k+1}} \equiv a \pmod{N}$ and so $\gcd(x_{n_{k+1}}, N) = 1$, proving the inductive step. \square

20. (China TST 2005) Integers a_0, a_1, \dots, a_n and x_0, x_1, \dots, x_n satisfy

$$a_0 x_0^k + a_1 x_1^k + \dots + a_n x_n^k = 0$$

for all $1 \leq k \leq r$, where r is a positive integer. Prove that m divides $a_0 x_0^m + a_1 x_1^m + \dots + a_n x_n^m$ for all $r + 1 \leq m \leq 2r + 1$.

Proof. Take $r + 1 \leq m \leq 2r + 1$ and let p be a prime factor of m and $u = v_p(m)$. It suffices to prove that $p^u \mid \sum_{j=0}^n a_j x_j^m$. We claim that $x_j^m \equiv x_j^{\frac{m}{p}} \pmod{p^u}$ for $0 \leq j \leq n$. If this holds, we obtain

$$\sum_{j=0}^n a_j x_j^m \equiv \sum_{j=0}^n a_j x_j^{\frac{m}{p}} \pmod{p^u}$$

and the last sum vanishes by assumption, since $\frac{m}{p} \in \{1, 2, \dots, r\}$. To prove the claim, we discuss two cases. If $p \mid x_j$, then $p^u \mid x_j^{\frac{m}{p}}$, since $\frac{m}{p} \geq p^{u-1} \geq u$. If p does not divide x_j then $\varphi(p^u) \mid m - \frac{m}{p}$ and so by

Euler's theorem $x_j^m \equiv x_j^{\frac{m}{p}} \pmod{p^u}$. This proves the claim and finishes the solution of the problem. \square

21. (Hong Kong 2010) Let n be an integer greater than 1 and let $1 \leq a_1 < \dots < a_k \leq n$ be the totatives of n . Prove that for any integer a relatively prime to n we have

$$\frac{a^{\phi(n)} - 1}{n} \equiv \sum_{i=1}^k \frac{1}{aa_i} \left\lfloor \frac{aa_i}{n} \right\rfloor \pmod{n}$$

Proof. Consider the Euclidean division $aa_i = q_i n + r_i$ of aa_i by n . Since a is relatively prime to n and (a_1, \dots, a_k) is a reduced residue system modulo n , so is (aa_1, \dots, aa_k) , thus r_1, \dots, r_k are just a permutation of a_1, \dots, a_k . In particular $\prod_{i=1}^k a_i = \prod_{i=1}^k r_i$.

Next, take the product of the relations $aa_i = q_i n + r_i$ and obtain

$$a^{\varphi(n)} a_1 \dots a_k = (q_1 n + r_1) \dots (q_k n + r_k).$$

Expanding the product in the right-hand side and reducing modulo n^2 yields

$$a^{\varphi(n)} a_1 \dots a_k \equiv r_1 \dots r_k + r_2 \dots r_k q_1 n + \dots + r_1 \dots r_{k-1} q_k n \pmod{n^2}.$$

Since $\prod_{i=1}^k a_i = \prod_{i=1}^k r_i$, we obtain the equivalent congruence

$$\frac{a^{\varphi(n)} - 1}{n} \equiv \frac{q_1}{r_1} + \dots + \frac{q_k}{r_k} \pmod{n}.$$

We conclude using the fact that $\frac{q_i}{r_i} \equiv \frac{q_i}{aa_i} \pmod{n}$ and that $q_i = \left\lfloor \frac{aa_i}{n} \right\rfloor$. \square

22. (Kömal) Let x_1, x_2, \dots, x_n be integers such that $\gcd(x_1, \dots, x_n) = 1$. Prove that if $s_i = x_1^i + x_2^i + \dots + x_n^i$, then

$$\gcd(s_1, s_2, \dots, s_n) \mid \text{lcm}(1, 2, \dots, n).$$

Proof. By example 6.7 it suffices to prove that $p^d \leq n$ for any prime p and any $d \geq 1$ such that $p^d \mid \gcd(s_1, \dots, s_n)$. Write

$$(X - x_1) \dots (X - x_n) = X^n + a_{n-1}X^{n-1} + \dots + a_0$$

for some integers a_0, \dots, a_n . Then $x_i^n + a_{n-1}x_i^{n-1} + \dots + a_0 = 0$ for $1 \leq i \leq n$. Multiplying this relation by x_i^r (for r an arbitrary nonnegative integer) and adding the resulting relations, we obtain

$$s_{n+r} + a_{n-1}s_{n+r-1} + \dots + a_0s_r = 0$$

for all $r \geq 0$. Since p^d divides s_1, \dots, s_n an immediate induction using the previous relation shows that p^d divides s_r for all $r \geq 1$. Setting $r = d \cdot \varphi(p^d)$ we obtain

$$x_1^{d\varphi(p^d)} + \dots + x_n^{d\varphi(p^d)} \equiv 0 \pmod{p^d}.$$

By Euler's theorem, each term $x_i^{d\varphi(p^d)}$ is either a multiple of p^d or congruent to 1 modulo p^d . Since $\gcd(x_1, \dots, x_n) = 1$, there is at least one nonzero term. We immediately conclude that $p^d \leq n$, as desired. \square

23. (Brazil 2005) Let a and c be positive integers. Prove that for any integer b there is a positive integer x such that

$$a^x + x \equiv b \pmod{c}.$$

Proof. The solution is very similar to that of example 7.55. We will prove by strong induction on c the following statement: for all integers b and all $a \geq 1$ there are infinitely many $x \geq 1$ such that $a^x + x \equiv b \pmod{c}$. The case $c = 1$ is clear, so assume that the result holds up to $c - 1$ and let us prove it for c . Fix $a \geq 1$ and an integer b . Since $\varphi(c) < c$, by the inductive hypothesis there are infinitely many $x \geq 1$ such that $\varphi(c) \mid a^x + x - b$. We can thus choose (once and for all) such x , with $x \geq \max_{p \mid c} v_p(c)$ and $x \geq b$. Then arguing as in example 7.55 we obtain $a^{x+k\varphi(c)} \equiv a^x \pmod{c}$ for all $k \geq 1$. Write $a^x + x - b = d\varphi(c)$ for

some $d \geq 1$ and set $y_k = x + (kc - d)\varphi(c)$ for $k > d$ ($k > d/c$ would be enough). Then

$$a^{y_k} + y_k \equiv a^x + y_k \equiv a^x + x - d\varphi(c) \equiv b \pmod{c},$$

thus all $(y_k)_{k>d}$ are solutions of the congruence $a^y + y \equiv b \pmod{c}$, proving the inductive step. \square

24. (Ibero American 2012) Prove that for any integer $n > 1$ there exist n consecutive positive integers such that none of them is divisible by the sum of its digits.

Proof. Write $s(x)$ for the sum of the digits of x . Choose pairwise distinct prime numbers p_1, p_2, \dots, p_n such that $p_i > \max(3, i)$ for $i \leq n$. Let $P = p_1 p_2 \dots p_n$ and consider

$$B = 10^{\varphi(P)} \frac{10^{(P-10)\varphi(P)} - 1}{10^{\varphi(P)} - 1} + 10.$$

Then B is a multiple of P by Euler's theorem, and $s(B) = P - 9$. Since $\gcd(P - 9, p_i) = 1$ for $1 \leq i \leq n$, the Chinese remainder theorem yields the existence of a positive integer t such that $t(P - 9) + s(i) \equiv 0 \pmod{p_i}$ for $1 \leq i \leq n$. Define $C = BB \dots B$ (with t copies of B) and note that for x large enough we have

$$s(10^x C + i) = s(C) + s(i) \equiv 0 \pmod{p_i}$$

for $i \leq n$. Then $10^x C + 1, \dots, 10^x C + n$ are consecutive numbers and none of them is a multiple of the sum of its digits. Indeed, if $s(10^x C + i)$ divides $10^x C + i$, then p_i divides $10^x C + i$ and since $p_i \mid C$ (because $p_i \mid B$) we deduce that $p_i \mid i$, a contradiction. \square

25. (Russia 2006) Let x and y be purely periodic decimal fractions such that $x + y$ and xy are purely periodic decimal fractions with period length T . Prove that the lengths of the periods of x and y are not greater than T .

Proof. Write $x = \frac{a}{b}$ and $y = \frac{c}{d}$ with $c, d > 0$ and $\gcd(a, b) = \gcd(c, d) = 1$. Write

$$x + y = \frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd} = \frac{e}{f}$$

and $xy = \frac{ac}{bd} = \frac{g}{h}$ in lowest form.

By assumption $f \mid 10^T - 1$ and $h \mid 10^T - 1$. Since $f(ad + bc) = ebd$, we obtain $b \mid fad$ and so $b \mid fd \mid d(10^T - 1)$. Next, since $ach = bdg$, we have $b \mid ach$ and so $b \mid ch \mid c(10^T - 1)$.

It follows that $b \mid \gcd(c(10^T - 1), d(10^T - 1)) = 10^T - 1$ and similarly we obtain $d \mid 10^T - 1$. The result follows.

Here is an alternative argument: by the hypotheses, the polynomial

$$P(X) = (10^T - 1)(X - x)(X - y)$$

has integer coefficients and roots x and y . By the rational root theorem, it follows that the denominators of x and y divide the leading coefficient $10^T - 1$ of $P(X)$. Thus their periods divide T . \square

26. (Iran 2013) Let p be an odd prime and let d be a positive divisor of $p - 1$. Let S be the set of integers $x \in \{1, 2, \dots, p - 1\}$ for which the order of x modulo p is d . Find the remainder of $\prod_{x \in S} x$ when divided by p .

Proof. For each $x \in S$ there is a unique $y \in \{1, 2, \dots, p - 1\}$ such that $xy \equiv 1 \pmod{p}$. If k is a positive integer, then the congruence $x^k \equiv 1 \pmod{p}$ is equivalent to $(xy)^k \equiv y^k \pmod{p}$ or $y^k \equiv 1 \pmod{p}$, thus the orders of x and y modulo p are the same and $y \in S$. Moreover, by construction $xy \equiv 1 \pmod{p}$. If $x = y$ then $x^2 \equiv 1 \pmod{p}$ and so $d = 1$ or $d = 2$. Thus for $d > 2$ the set S has a partition into pairs (x, y) as above and so $\prod_{x \in S} x \equiv 1 \pmod{p}$. If $d = 1$ then $S = \{1\}$ and the answer is again 1. If $d = 2$ then $S = \{1, p - 1\}$ and the answer is -1 . \square

27. Let a, b, n be positive integers with $a \neq b$. Prove that

$$2n \mid \varphi(a^n + b^n) \quad \text{and} \quad n \mid \varphi\left(\frac{a^n - b^n}{a - b}\right).$$

Proof. Note that we may assume that $\gcd(a, b) = 1$, by replacing a and b with $\frac{a}{\gcd(a, b)}$ and $\frac{b}{\gcd(a, b)}$. We may also assume that $a > b$, by symmetry.

We prove first the divisibility $2n \mid \varphi(a^n + b^n)$. Let c be a positive integer such that $bc \equiv 1 \pmod{a^n + b^n}$ (c exists since $\gcd(b, a^n + b^n) = 1$) and note that $\gcd(ac, a^n + b^n) = 1$. If $k \geq 1$ the congruence $(ac)^k \equiv 1 \pmod{a^n + b^n}$ is equivalent to $(abc)^k \equiv b^k \pmod{a^n + b^n}$ and then to $a^k \equiv b^k \pmod{a^n + b^n}$. Let d be the order of ac modulo $a^n + b^n$. Since $a^{2n} \equiv b^{2n} \pmod{a^n + b^n}$, the previous discussion yields $d \mid 2n$ and $a^n + b^n \mid a^d - b^d$, in particular $d \mid 2n$ and $d > n$. We deduce that $d = 2n$. Since $d \mid \varphi(a^n + b^n)$, the result follows.

We prove next the divisibility $n \mid \varphi\left(\frac{a^n - b^n}{a - b}\right)$. Let

$$N = \frac{a^n - b^n}{a - b} = a^{n-1} + a^{n-2}b + \dots + b^{n-1}.$$

Since $\gcd(ab, N) = 1$, there is a positive integer c such that $bc \equiv 1 \pmod{N}$, and we have $\gcd(ac, N) = 1$. Arguing as above, one proves that the order of ac modulo N equals n and concludes that $n \mid \varphi(N)$. \square

28. Find all primes p and q such that $p^2 + 1 \mid 2003^q + 1$ and $q^2 + 1 \mid 2003^p + 1$.

Proof. We may assume that $p \leq q$. If $p = 2$ then $5 \mid 2003^q + 1$, which forces $q = 2$ and gives the solution $(2, 2)$. Assume next that $p > 2$. If r is a prime factor of $p^2 + 1$, then $r \mid 2003^{2q} - 1$, hence $\text{ord}_r(2003) \mid 2q$. Assuming that $\gcd(q, \text{ord}_r(2003)) = 1$, we infer that

$$r \mid 2003^2 - 1 = 2^3 \cdot 3 \cdot 7 \cdot 11 \cdot 13 \cdot 167.$$

Since $r \equiv 1 \pmod{4}$ (combine corollary 5.28 with $r \mid p^2 + 1$), it follows that $r = 2$ or $r = 13$. We cannot have $r = 13$ since $r \mid 2003^q + 1$ and $2003^q + 1 \equiv 2 \pmod{13}$. We conclude that $r = 2$ or $q \mid \text{ord}_r(2003) \mid r - 1$, i.e. any odd prime factor r of $p^2 + 1$ is congruent to 1 modulo q . Combined with $v_2(p^2 + 1) = 1$, this gives $p^2 + 1 \equiv 2 \pmod{q}$, i.e. $q \mid (p - 1)(p + 1)$. This is however impossible since $q \geq p$ and p is odd. Thus the only solution of the problem is $(2, 2)$. \square

29. (MOSP 2001) Let p be a prime number and let m, n be integers greater than 1 such that $n \nmid m^{p(n-1)} - 1$. Prove that $\gcd(m^{n-1} - 1, n) > 1$.

Proof. Assume that $\gcd(m^{n-1} - 1, n) = 1$. Let $a = v_p(n - 1)$, let q be any prime factor of n and finally let d be the order of $m \bmod q$. Since q does not divide $m^{n-1} - 1$, d cannot divide $n - 1$. On the other hand, q divides $m^{p(n-1)} - 1$, thus d divides $p(n - 1)$. It follows that $v_p(d) > a$ and since $d \mid q - 1$, we obtain $v_p(q - 1) \geq a + 1$. Since this happens for all primes q dividing n , it follows that $n \equiv 1 \pmod{p^{a+1}}$, which contradicts the equality $v_p(n - 1) = a$. The result follows. \square

30. a) (Pepin's test) Let n be a positive integer and let $k = 2^{2^n} + 1$. Prove that k is a prime if and only if $k \mid 3^{\frac{k-1}{2}} + 1$.
- b) (Euler-Lagrange) Let $p \equiv -1 \pmod{4}$ be a prime. Prove that $2p + 1$ is a prime if and only if $2p + 1 \mid 2^p - 1$.

Proof. a) Suppose that k is a prime and let us prove that $k \mid 3^{\frac{k-1}{2}} + 1$. By Euler's criterion (theorem 5.99) this is equivalent to $\left(\frac{3}{k}\right) = -1$ and, by the quadratic reciprocity law (theorem 5.124), to $(-1)^{\frac{k-1}{2}} \cdot \left(\frac{k}{3}\right) = -1$. This follows directly from $k \equiv 1 \pmod{4}$ and $k \equiv 2 \pmod{3}$.

Suppose next that $k \mid 3^{\frac{k-1}{2}} + 1$ and let p be a prime factor of k . Since $p \mid k \mid 3^{k-1} - 1$, the order d of 3 modulo p divides $k - 1 = 2^{2^n}$, so d is a power of 2. If $d < k - 1$, then $d \mid \frac{k-1}{2}$ and $p \mid 3^{\frac{k-1}{2}} - 1$, contradicting the fact that $p \mid k \mid 3^{\frac{k-1}{2}} + 1$. Thus $d = k - 1$ and since $d \mid p - 1$, we obtain $k \leq p$. Since $p \mid k$, we conclude that $p = k$, as needed.

b) The argument is very similar. If $q = 2p + 1$ is a prime, we need to prove that $q \mid 2^{\frac{q-1}{2}} - 1$, i.e. that $\left(\frac{2}{q}\right) = 1$, which follows from the fact that $q \equiv -1 \pmod{8}$ and theorem 5.125. Conversely, if $q \mid 2^p - 1$, then the order of 2 mod q must be p (since it is not 1 and it divides p), and the same happens for any prime factor l of q . Thus $p \mid l - 1$ for each prime factor l of q , which immediately yields the result. \square

Remark 8.27. Pepin's test was used for instance to prove that each of the numbers $F_{13}, F_{14}, F_{20}, F_{22}, F_{24}$ is composite, where $F_n = 2^{2^n} + 1$.

31. Let $p > 2$ be an odd prime and let a be a primitive root modulo p . Prove that $a^{\frac{p-1}{2}} \equiv -1 \pmod{p}$.

Proof. By Fermat's little theorem we have $a^{p-1} \equiv 1 \pmod{p}$, hence $(a^{\frac{p-1}{2}} - 1)(a^{\frac{p-1}{2}} + 1) \equiv 0 \pmod{p}$. We cannot have $a^{\frac{p-1}{2}} \equiv 1 \pmod{p}$, since the order of $a \pmod{p}$ is $p-1 > \frac{p-1}{2}$. Hence $a^{\frac{p-1}{2}} + 1 \equiv 0 \pmod{p}$, as desired. \square

32. Suppose that $n > 1$ is an integer for which there are primitive roots modulo n . Prove that the set $\{1, 2, \dots, n\}$ contains exactly $\varphi(\varphi(n))$ primitive roots modulo n .

Proof. Pick a primitive root g modulo n . If a is another primitive root modulo n , then we can find an integer $k \in \{0, 1, \dots, \varphi(n) - 1\}$ such that $a \equiv g^k \pmod{n}$.

Since the order of g modulo n is $\varphi(n)$, proposition 7.66 shows that the order of g^k modulo n is $\varphi(n)$ if and only if $\gcd(k, \varphi(n)) = 1$. We deduce that the primitive roots modulo n in the set $\{1, 2, \dots, n\}$ are precisely the remainders modulo n of the numbers g^{a_1}, \dots, g^{a_k} , where a_1, \dots, a_k are the totatives of $\varphi(n)$. Their number is $k = \varphi(\varphi(n))$. \square

33. Let p be an odd prime. Prove that p is a Fermat prime (i.e. of the form $2^n + 1$ with $n \geq 1$) if and only if every quadratic non-residue mod p is a primitive root mod p .

Proof. There are $\frac{p-1}{2}$ quadratic non-residues mod p and $\varphi(p-1)$ primitive roots mod p (use the previous exercise or Theorem 7.104 for the last assertion). Thus if any quadratic non-residue mod p is a primitive root mod p we must have $\varphi(p-1) \geq \frac{p-1}{2}$. Writing $p-1 = 2^k m$ with $k \geq 1$ and m odd, the previous inequality becomes $\varphi(m) \geq m$ and forces $m = 1$. Thus p is a Fermat prime.

Conversely, assume that p is a Fermat prime, then $\varphi(p-1) = \frac{p-1}{2}$ since $p-1$ is a power of 2. Since any primitive root mod p is a quadratic non-residue mod p and since there are as many quadratic non-residues mod p as primitive roots mod p , it follows that every quadratic non-residue mod p is a primitive root mod p . The result follows. \square

34. Let $\lambda(n)$ be the least positive integer k such that $x^k \equiv 1 \pmod{n}$ for all x relatively prime to n . Prove that
- If k is a positive integer such that $x^k \equiv 1 \pmod{n}$ for all x relatively prime to n , then k is a multiple of $\lambda(n)$.
 - $\lambda(mn) = \text{lcm}(\lambda(m), \lambda(n))$ for m, n relatively prime.
 - We have $\lambda(n) = \varphi(n)$ when $n = 2, 4$ or a power of an odd prime, and $\lambda(2^n) = 2^{n-2}$ for $n \geq 3$.
 - For each n , the set of numbers $\text{ord}_n(x)$ (over all x relatively prime to n) is precisely the set of positive divisors of $\lambda(n)$.

Proof. a) Write $k = q\lambda(n) + r$ with $0 \leq r < \lambda(n)$. Assume that $r > 0$. Then for all x relatively prime to n we have

$$1 \equiv x^k = (x^{\lambda(n)})^q \cdot x^r \equiv x^r \pmod{n}.$$

This contradicts the minimality of $\lambda(n)$.

b) Let $M = \text{lcm}(\lambda(m), \lambda(n))$. Suppose that x is relatively prime to mn , so it is relatively prime to both m and n . Now by definition of M we have $x^M \equiv 1 \pmod{n}$ and $x^M \equiv 1 \pmod{m}$, thus $x^M \equiv 1 \pmod{mn}$, since $\gcd(m, n) = 1$. Since x was arbitrary, this yields (thanks to a)) $\lambda(mn) \mid M$.

To prove that $M \mid \lambda(mn)$, it suffices, thanks to a) and to symmetry in m and n , to prove that $x^{\lambda(mn)} \equiv 1 \pmod{n}$ for all x relatively prime to n . Take such x . Note that x is not necessarily prime to m , but since $\gcd(m, n) = 1$ we can find y such that $y \equiv x \pmod{n}$ and $y \equiv 1 \pmod{m}$ (Chinese remainder theorem). Now y is relatively prime to mn , so $y^{\lambda(mn)} \equiv 1 \pmod{n}$. But clearly $y^{\lambda(mn)} \equiv x^{\lambda(mn)} \pmod{n}$, hence the result.

c) Note that in all cases $\lambda(n) \mid \varphi(n)$, thanks to a) and Euler's theorem.

Suppose that n is 2, 4 or a power of an odd prime. Then we can find a primitive root g modulo n . Since $g^{\lambda(n)} \equiv 1 \pmod{n}$ and since g has order $\varphi(n)$ modulo n , it follows that $\varphi(n) \mid \lambda(n)$. Combined with the first paragraph, this yields $\varphi(n) = \lambda(n)$ for such n .

Now consider 2^n for $n \geq 3$. By the first paragraph $\lambda(2^n)$ divides $\varphi(2^n) = 2^{n-1}$, hence it is a power of 2, say 2^k . Thus we need to find the smallest k for which $x^{2^k} \equiv 1 \pmod{2^n}$ for all odd numbers x . We saw several times that $k = n - 2$.

d) Let x be an integer relatively prime to n and let $d = \text{ord}_n(x)$. Since $x^{\lambda(n)} \equiv 1 \pmod{n}$, we deduce that $d \mid \lambda(n)$. Conversely, let d be a positive divisor of $\lambda(n)$. We need to prove that we can find x relatively prime to n such that $\text{ord}_n(x) = d$.

Let $n = p_1^{k_1} \dots p_s^{k_s}$. By b) and c) we have $\lambda(n) = \text{lcm}(\varphi(p_1^{k_1}), \dots, \varphi(p_s^{k_s}))$. Let $d_i = \gcd(d, \varphi(p_i^{k_i}))$. Since every prime power factor of d must divide at least one of the $\varphi(p_i^{k_i})$, we have $d = \text{lcm}(d_1, \dots, d_s)$. Let a_i be a primitive root modulo $p_i^{k_i}$. Then $x_i = a_i^{\varphi(p_i^{k_i})/d_i}$ has order d_i modulo $p_i^{k_i}$. Using the Chinese remainder theorem, we can find an integer x such that $x \equiv x_i \pmod{p_i^{k_i}}$ for all i . Then x is clearly relatively prime to n , and its order modulo $p_i^{k_i}$ is d_i . Hence its order modulo n is a multiple of $d = \text{lcm}(d_1, \dots, d_s)$. Conversely, d_i divides d for all i and hence $x^d \equiv 1 \pmod{p_i^{k_i}}$ for all i . Thus $x^d \equiv 1 \pmod{n}$ and hence the order of x modulo n divides d . Thus the order must be d . \square

35. Let $p > 2$ be a prime and let a be a primitive root mod p . Prove that $-a$ is a primitive root mod p if and only if $p \equiv 1 \pmod{4}$.

Proof. Note that if x is a primitive root modulo p then $x^{\frac{p-1}{2}} \equiv -1 \pmod{p}$, since $(x^{\frac{p-1}{2}})^2 \equiv 1 \pmod{p}$ by Fermat's little theorem and $x^{\frac{p-1}{2}}$ is not congruent to 1 modulo p since $\text{ord}_p(x) = p - 1$. Suppose that $-a$ is a primitive root modulo p , then by the preliminary discussion we have $a^{\frac{p-1}{2}} \equiv (-a)^{\frac{p-1}{2}} \equiv -1 \pmod{p}$, thus $(-1)^{\frac{p-1}{2}} \equiv 1 \pmod{p}$

and since $p > 2$ we deduce that $p \equiv 1 \pmod{4}$. Conversely, suppose that $p \equiv 1 \pmod{4}$ and let d be the order of $-a$ modulo p . Then $(-a)^d \equiv 1 \pmod{p}$, thus $a^{2d} \equiv 1 \pmod{p}$ and since a is a primitive root modulo p we deduce that $p-1 \mid 2d$. If $-a$ is not a primitive root modulo p , then necessarily $d = \frac{p-1}{2}$ and so $(-a)^{\frac{p-1}{2}} \equiv 1 \pmod{p}$. But $(-a)^{\frac{p-1}{2}} \equiv (-1)^{\frac{p-1}{2}} a^{\frac{p-1}{2}} \equiv 1 \cdot (-1) \equiv -1 \pmod{p}$, a contradiction. Thus $-a$ is a primitive root mod p and the problem is solved. \square

36. (Unesco Competition 1995) Let m, n be integers greater than 1. Prove that the remainders of the numbers $1^n, 2^n, \dots, m^n$ modulo m are pairwise distinct if and only if m is square-free and n is relatively prime to $\varphi(m)$.

Proof. Suppose that the remainders of $1^n, 2^n, \dots, m^n$ are pairwise distinct. If there is a prime p such that $p^2 \mid m$, then m^n and $(m/p)^n$ are both 0 modulo m , a contradiction. Thus m is squarefree, say $m = p_1 \dots p_k$ for some pairwise distinct primes p_1, \dots, p_k . We need to prove that n is relatively prime to $p_i - 1$ for all i , since $\varphi(m) = (p_1 - 1) \dots (p_k - 1)$. Suppose that $d_i = \gcd(p_i - 1, n) > 1$ for some i , and pick a primitive root g modulo p_i . Then $x = g^{\frac{p_i-1}{d_i}}$ satisfies $x^n \equiv 1 \pmod{p_i}$ and x is not congruent to 1 modulo p_i . Thanks to the Chinese remainder theorem we can find y such that $y \equiv 1 \pmod{p_j}$ for all $j \neq i$ and $y \equiv x \pmod{p_i}$. But then $y^n \equiv 1 \pmod{m}$ and y is not congruent to 1 modulo m , a contradiction. This proves one direction.

Next, assume that $m = p_1 \dots p_k$ is squarefree and $\gcd(n, \varphi(m)) = 1$. Suppose that for some $1 \leq i < j \leq m$ we have $i^n \equiv j^n \pmod{m}$. Then $i^n \equiv j^n \pmod{p_r}$ for all r and since $p_r - 1$ and n are relatively prime, we deduce that $i \equiv j \pmod{p_r}$. But then $i \equiv j \pmod{m}$, which is impossible. This proves the opposite direction and finishes the solution. \square

37. (adapted from Tuymaada 2011) Prove that among 2500 consecutive positive integers there is an integer n such that the length of the period of the decimal expansion of $\frac{1}{n}$ is greater than 2011.

Proof. We want to ensure that the order of 10 modulo n is greater than 2011. It suffices to prove that there is $d \in \{1, 2, \dots, 2500\}$ such that the order of 10 modulo d is greater than 2011. Indeed, if such d exists, then among any 2500 consecutive integers one can find a multiple of d , and the result follows.

We start by finding a prime p for which 10 is a primitive root modulo p . Trial and error gives the smallest answer $p = 7$. Let us see whether 10 is a primitive root modulo 49. The order of 10 modulo 49 is a multiple of 6 and a divisor of $\varphi(49) = 42$. So if 10 is not a primitive root modulo 49, then its order is 6 and so $49 \mid 10^6 - 1 = 3^3 \cdot 7 \cdot 11 \cdot 13 \cdot 37$. The last result is absurd, hence 10 is a primitive root modulo 49, thus also a primitive root modulo 7^n for all n , by theorem 7.108. Now, it is enough to ensure that we can find a power of 7 with $\varphi(7^k) > 2011$ and $7^k \leq 2500$. It is not difficult to see that 7^4 works, hence the result. \square

38. Is there a positive integer which is divisible by the product P of its digits and such that P is a power of 7 greater than 10^{2016} ?

Proof. Fix a positive integer k . We will prove that there is a number x divisible by the product P of its digits and such that $P = 7^k$. We saw in the previous exercise that 10 is a primitive root modulo 7^k . Consider the number $a = \overline{66 \dots 6} = 6 \cdot \frac{10^k - 1}{9}$, then $1 - 9a = 7 - 6 \cdot 10^k$ is relatively prime to 7 and so there are infinitely many n such that $10^n \equiv 1 - 9a \pmod{7^k}$. For such n which is greater than k , the number $\frac{10^n - 1}{9} + a$ has the product of its digits equal to 7^k and is a multiple of 7^k by construction. \square

39. Let m, n be positive integers. Prove that there is a positive integer k such that $2^k \equiv 1999 \pmod{3^m}$ and $2^k \equiv 2009 \pmod{5^n}$.

Proof. Using theorem 7.108, it follows easily that 2 is a primitive root modulo 3^m and modulo 5^n , thus we can find positive integers k_1, k_2 such that $2^{k_1} \equiv 1999 \pmod{3^m}$ and $2^{k_2} \equiv 2009 \pmod{5^n}$. Considering these simply as congruences modulo 3 and 5, it follows that k_1 and k_2 are even. The Chinese remainder theorem yields the existence of a positive integer

a such that $a \equiv \frac{k_1}{2} \pmod{3^{m-1}}$ and $a \equiv \frac{k_2}{2} \pmod{2 \cdot 5^{n-1}}$. Letting $k = 2a$ and using Euler's theorem, we obtain $2^k \equiv 1999 \pmod{3^m}$ and $2^k \equiv 2009 \pmod{5^n}$, as desired. \square

40. (Iran 2012) Let p be an odd prime. Prove that there is a positive integer x such that x and $4x$ are both primitive roots modulo p .

Proof. Fix a primitive root g modulo p and write $2 \equiv g^k \pmod{p}$ for some positive integer k . If we can find a positive integer a such that a and $a + 2k$ are relatively prime to $p - 1$ then $x = g^a$ and $4x \equiv g^{a+2k} \pmod{p}$ are primitive roots modulo p (see the solution of exercise 32) and we are done. Let q_1, \dots, q_s be the prime factors of $p - 1$, with $q_1 = 2$. If $i > 1$ then $q_i > 2$ hence there is $a_i \in \{0, 1, \dots, q_i - 1\}$ different from 0 and $-2k \pmod{q_i}$. Set $a_1 = 1$. By the Chinese remainder theorem we can choose an integer a such that $a \equiv a_i \pmod{q_i}$ for all $1 \leq i \leq s$. Then by construction a and $a + 2k$ are not divisible by q_i for any i , so they are relatively prime to $p - 1$. The result follows. \square

41. (Brazil 2009) Let p, q be odd primes such that $q = 2p + 1$. Prove that there is a multiple of q whose sum of digits is 1, 2 or 3.

Proof. The order d of 10 modulo q divides $q - 1 = 2p$, so it equals 1, 2, p or $2p$. If $d = 1$ then q divides 9 and hence $q = 3$, which is not possible. If $d = 2$, then q divides 99 and hence $q = 11$ and we can choose 11 to be the desired multiple. If $d = 2p$, then $q \mid 10^p + 1$, and we can choose $10^p + 1$ as a multiple of q with sum of digits 2. This leaves only the case $d = p$. In this case $10^p = 10^{(q-1)/2} \equiv 1 \pmod{q}$ and hence 10 is a quadratic residue modulo q . Hence any power of 10 is also a quadratic residue modulo q . Since there are p distinct powers of 10 modulo q and there are p nonzero quadratic residues modulo q , we deduce that every nonzero quadratic residue modulo q is a power of 10. Thus if we can find integers x, y that are relatively prime to q and such that $q \mid x^2 + y^2 + 1$ we can find integers a and b such that $q \mid 10^a + 10^b + 1$ which will give the desired multiple. Consider the set A of all quadratic residues modulo

q , and the set B consisting of the remainders modulo q of the numbers $-1 - a$ with $a \in A$. Then A, B are subsets of $\{0, 1, \dots, q-1\}$, each with $\frac{q+1}{2}$ elements, thus $A \cap B \neq \emptyset$. It follows that there are integers x, y such that $q \mid x^2 + y^2 + 1$. Since $q \equiv 3 \pmod{4}$ (which follows from $q = 2p + 1$ and the fact that p is odd), corollary 5.28 shows that x and y are necessarily prime to q , which finishes the proof. \square

42. (Brazil 2012) Find the least positive integer n for which there is a positive integer k such that the last 2012 decimal digits of n^k are all 1's.

Proof. Let $s = \frac{10^{2012}-1}{9}$. We need to understand the congruence $n^k \equiv s \pmod{10^{2012}}$. Note that $9s \equiv -1 \pmod{10^{2012}}$ and in particular $9s \equiv -1 \pmod{16}$ and $9s \equiv -1 \pmod{5}$. Thus we find $s \equiv 7 \pmod{16}$ and $s \equiv 1 \pmod{5}$. Hence n is clearly odd and k is clearly odd (if k is even, then $n^k \equiv 1 \pmod{8}$). Hence $n^k \equiv n \pmod{8}$, and so $n \equiv 7 \pmod{8}$. If $n \equiv 15 \pmod{16}$, then we would get $n^2 \equiv 1 \pmod{16}$ and $n^k \equiv 15 \pmod{16}$. Thus $n \equiv 7 \pmod{16}$. The order of n modulo 5 must divide both 4 and k , and since k is odd, the order must be 1, i.e. $n \equiv 1 \pmod{5}$. Combining these congruences, we obtain $n \equiv 71 \pmod{80}$ and so $n \geq 71$.

We will embark now on the technical part of the proof, showing that $n = 71$ is a solution of the problem, i.e. proving the existence of $k \geq 1$ such that $9 \cdot 71^k \equiv -1 \pmod{10^{2012}}$. First, proposition 7.66 gives $\text{ord}_{2^N}(71) = 2^{N-3}$ and $\text{ord}_{5^N}(71) = 5^{N-1}$ for all $N \geq 2$. Next, we will prove in the next paragraph that for all $N \geq 4$ we can find $l \geq 1$ such that $9 \cdot 71^l \equiv -1 \pmod{2^N}$ and $m \geq 1$ such that $9 \cdot 71^m \equiv -1 \pmod{5^N}$. Assuming this for a moment, we can use the Chinese remainder theorem to obtain the existence of an integer $k \geq 1$ such that $k \equiv l \pmod{2^{N-3}}$ and $k \equiv m \pmod{5^{N-1}}$. Then $9 \cdot 71^k \equiv -1 \pmod{10^N}$ and taking $N = 2012$ finishes the proof.

It remains to prove the existence of l and m . Since $\text{ord}_{2^N}(71) = 2^{N-3}$, we deduce that the numbers $1, 71, \dots, 71^{2^{N-3}-1}$ give 2^{N-3} pairwise distinct remainders modulo 2^N , and since $71 \equiv 7 \pmod{16}$ and $71^2 \equiv 1 \pmod{16}$, all these remainders are of the form $7 + 16r$ or $1 + 16r$. On the

other hand, there are exactly 2^{N-4} remainders modulo 2^N of the form $7 + 16r$ and 2^{N-4} remainders of the form $1 + 16r$. We deduce that the remainders of $1, 71, \dots, 71^{2^{N-3}-1}$ modulo 2^N are precisely all remainders of the form $7 + 16r$ and $1 + 16r$. Since $s \equiv 7 \pmod{16}$, it must be one of these remainders and hence l exists. Similarly, the powers of 71 modulo 5 are exactly the numbers congruent to 1 modulo 5, and hence m exists. In conclusion the answer is $n = 71$. \square

43. (Nieuw Archief voor Wiskunde) Suppose that $\alpha \geq \frac{\log 10}{\log 5} = 1.43067\dots$. Prove that for any $n \geq 1$, any sequence of n digits (between 0 and 9) occurs as a sequence of consecutive digits in the last $\lceil \alpha n \rceil$ digits of some power of 2.

Proof. Consider $a_0, \dots, a_{n-1} \in \{0, 1, \dots, 9\}$ and let

$$A = a_0 + 10a_1 + \dots + 10^{n-1}a_{n-1} = \overline{a_{n-1}\dots a_0}.$$

We will prove that we can find $\lceil \alpha n \rceil - n \geq s \geq 1$ digits $\alpha_1, \dots, \alpha_s$ and $k \geq n + s$ such that setting

$$B = \overline{A\alpha_1\dots\alpha_s} = A \cdot 10^s + \alpha_1 \cdot 10^{s-1} + \dots + \alpha_s$$

we have $2^k \equiv B \pmod{10^{n+s}}$. If we succeed in proving this, then the last $n+s$ digits of 2^k are the digits of B and since $s \leq \lceil \alpha n \rceil - n$, we deduce that A occurs among the last $\lceil \alpha n \rceil$ digits of 2^k . Let $r = \lceil \alpha n \rceil - n$ and observe that our hypothesis yields $5^{\lceil \alpha n \rceil} > 10^n$, thus $2^{n+r} < 10^r$. It follows that there is a multiple Y of 2^{n+r} in the interval $(A \cdot 10^r, (A+1) \cdot 10^r)$. Let s be the smallest positive integer for which we can find a multiple B of 2^{n+s} in $(A \cdot 10^s, (A+1) \cdot 10^s)$. We have just established that $s \leq r = \lceil \alpha n \rceil - n$. We claim that 5 does not divide B . Indeed, otherwise $10 \mid B$ and so $s \geq 2$, but then $\frac{B}{10} \in (A \cdot 10^{s-1}, (A+1) \cdot 10^{s-1})$ and $2^{n+s-1} \mid \frac{B}{10}$, a contradiction with the minimality of s . Since 2 is a primitive root modulo 5^n (use theorem 7.108), there is an integer $k \geq n + s$ such that $B \equiv 2^k \pmod{5^{n+s}}$ and then clearly $B \equiv 2^k \pmod{10^{n+s}}$, as desired. \square

Remark 8.28. One can prove that $\frac{\log 10}{\log 5}$ is the least number with this property.

44. Find all sequences of positive integers $(a_n)_{n \geq 1}$ such that

- a) $m - n \mid a_m - a_n$ for all positive integers m, n ;
- b) If m, n are relatively prime, then a_m and a_n are relatively prime.

Proof. We prove first that a_p is a power of p for any prime p . Suppose that $p \neq q$ are primes and $q \mid a_p$. Since $q \mid a_{p+q} - a_p$, we have $q \mid a_{p+q}$, but then $\gcd(a_p, a_{p+q}) > 1$ while $\gcd(p, p+q) = 1$, a contradiction. Thus we can find a sequence of nonnegative integers $(n_p)_p$ indexed by the set of odd primes such that $a_p = p^{n_p}$ for all p .

Next, we prove that $n_p = n_q$ for all odd primes p, q . Fix a positive integer m and set $u = \frac{p^{2^m} + 1}{2}$, an odd number since 4 does not divide $p^{2^m} + 1$. Choose an integer $n \geq m$ such that $v := \frac{q^{2^n} + 1}{2}$ is relatively prime to u (such n exists, since the numbers $(\frac{q^{2^n} + 1}{2})_{n \geq 1}$ are pairwise relatively prime, by an argument identical to that in example 3.12). Combining the Chinese remainder theorem and Dirichlet's theorem yields the existence of a prime r such that

$$r \equiv p \pmod{u} \quad \text{and} \quad r \equiv q \pmod{v}.$$

We deduce that $u \mid r - p \mid a_r - a_p = r^{n_r} - p^{n_p}$ and also $u \mid r - p \mid r^{n_r} - p^{n_r}$, thus $u \mid p^{n_p} - p^{n_r}$. Since the order of p modulo u is 2^{m+1} (this follows using the standard argument starting with $p^{2^m} \equiv -1 \pmod{u}$) we deduce that $2^{m+1} \mid n_p - n_r$. A similar argument shows that $2^{n+1} \mid n_q - n_r$ and since $n \geq m$ we deduce that $2^{m+1} \mid n_p - n_q$. Since m was arbitrary, it follows that $n_p = n_q$, as claimed.

Write n for the common value $n_3 = n_5 = n_7 = n_{11} = \dots$, so that $a_p = p^n$ for all odd primes p . If k is a positive integer and p is an odd prime, then $p - k \mid a_p - a_k = p^n - a_k$ and since $p - k \mid p^n - k^n$, we deduce that $p - k \mid a_k - k^n$. Since this happens for any odd prime p , we obtain $a_k = k^n$ for all $k \geq 1$. Conversely, it is clear that for any $n \geq 0$ the sequence $(a_k = k^n)_{k \geq 1}$ satisfies the desired properties. \square

45. (adapted after China TST 2012) Let $n > 1$ be an integer. Find all functions $f : \mathbf{Z} \rightarrow \{1, 2, \dots, n\}$ such that for each $k \in \{1, 2, \dots, n-1\}$ there is $j(k) \in \mathbf{Z}$ such that for all integers m we have

$$f(m + j(k)) \equiv f(m + k) - f(m) \pmod{n+1}.$$

Proof. Let f be a solution of the problem. Since f takes values between 1 and n the given congruence shows that $f(m + k) \neq f(m)$ for all m and all $1 \leq k \leq n-1$. It follows that for all $m \in \mathbf{Z}$ the numbers $f(m), f(m+1), \dots, f(m+n-1)$ are pairwise distinct and so they must be a permutation of $1, 2, \dots, n$. Applying this to m and $m+1$ we deduce that $f(m+n) = f(m)$ for all m , thus f is n -periodic. In particular, we may assume that $j(k) \in \{0, 1, \dots, n-1\}$ for all $1 \leq k \leq n-1$.

Note that if f is a solution of the problem, then so is $x \mapsto f(x+a)$ for any integer a . In particular, we may assume that $f(0) = 1$, since we know that f takes the value 1 by the first paragraph.

Then $f(j(k)) \equiv f(k) - 1 \pmod{n+1}$ for $1 \leq k \leq n-1$. If $j(k) \neq 0$, we deduce that $f(j(j(k))) \equiv f(k) - 2 \pmod{n+1}$. Similarly, if $j(j(k)) \neq 0$, then $f(j(j(j(k)))) \equiv f(k) - 3 \pmod{n+1}$. Since f does not take values which are multiples of $n+1$, we deduce that there is a smallest positive integer $a_k \leq f(k) \leq n$ such that $j^{a_k}(k) = 0$, where j^r denotes the r -fold composition of j with itself. Moreover, $1 = f(j^{a_k}(k)) \equiv f(k) - a_k \pmod{n+1}$. Since $f(1), \dots, f(n-1)$ are pairwise incongruent modulo $n+1$, we deduce that a_1, \dots, a_{n-1} are also pairwise incongruent modulo $n+1$, in particular pairwise distinct. Moreover, $a_1, \dots, a_{n-1} \in \{1, 2, \dots, n\}$ and we actually have $a_k \neq n$ for all k since if $a_k = n$ then the congruence $f(k) - a_k \equiv 1 \pmod{n}$ cannot be satisfied with $f(k) \in \{1, 2, \dots, n\}$. Hence a_1, \dots, a_{n-1} must be a permutation of $1, 2, \dots, n-1$.

Next, note that iterating the congruence $f(m + j(k)) \equiv f(m + k) - f(m) \pmod{n+1}$ we obtain $f(m + j^s(k)) \equiv f(m + k) - sf(m) \pmod{n+1}$ for $1 \leq s \leq a_k$, thus for $s = a_k$ we obtain $f(m + k) \equiv (a_k + 1)f(m) \pmod{n+1}$. But then $f(m + sk) \equiv (a_k + 1)^s f(m) \pmod{n+1}$ for all $s \geq 1$. Applying this to $s = n$ and using that f is n -periodic, we obtain $f(m)((a_k + 1)^n - 1) \equiv 0 \pmod{n+1}$ for all $m \in \mathbf{Z}$ and all $1 \leq k \leq n-1$.

Taking $m = 0$ we deduce that $n + 1 \mid (a_k + 1)^n - 1$ for $1 \leq k \leq n - 1$ and since a_1, \dots, a_{n-1} are a permutation of $1, 2, \dots, n - 1$, we deduce that $n + 1 \mid 2^n - 1, 3^n - 1, \dots, n^n - 1$. It follows that $n + 1$ is relatively prime to $1, 2, \dots, n$ and so $n + 1$ is a prime. In particular, if $n + 1$ is composite then there is no such function f .

Finally, suppose that $n + 1$ is a prime. Since $f(m + 1) \equiv (a_1 + 1)f(m) \pmod{n + 1}$, it follows that $f(m) \equiv (a_1 + 1)^m \pmod{n + 1}$ for $m \in \mathbf{Z}$. Since $f(0), f(1), \dots, f(n - 1)$ are a permutation of $1, 2, \dots, n$, we deduce that $g := a_1 + 1$ is a primitive root modulo $n + 1$ and $f(m)$ is the remainder of g^m modulo $n + 1$. Recalling that if f is a solution, then $x \mapsto f(a + x)$ is a solution for all integers a , we deduce that all solutions of the problem (when $n + 1$ is prime) are of the form $f(x) = ag^x \pmod{n + 1}$ for some $a \in \{1, 2, \dots, n\}$ and some primitive root g modulo $n + 1$. Conversely, it is easy to see that these are indeed solutions of the problem. \square

Bibliography

- [1] V. Boju, L. Funar, *The Math Problems Notebook*, Birkhauser, 2007.
- [2] Z. I. Borevich, I. R. Shafarevich, *Number Theory*, Academic Press (New York), 1966.
- [3] H. Davenport, *Multiplicative Number Theory*, 2nd ed., Springer-Verlag (New York), 1980.
- [4] T. Andreescu, G. Dospinescu, *Problems from the Book*, XYZ Press, 2008.
- [5] C. F. Gauss, *Disquisitiones Arithmeticae (Discourses on Arithmetic)*, English ed., Yale University Press (New Haven), 1966.
- [6] G. H. Hardy, E. M. Wright, *An Introduction to the Theory of Numbers*, 5th ed., Clarendon Press (Oxford), 1979.
- [7] K. Ireland, M. Rosen, *A Classical Introduction to Modern Number Theory*, Springer-Verlag (New York), 1982.
- [8] E. Landau, *Elementary Number Theory*, Chelsea (New York), 1958.
- [9] T. Nagell, *Introduction to Number Theory*, Chelsea (New York), 1981.
- [10] I. Niven, H. S. Zuckerman, H. L. Montgomery, *An Introduction to the Theory of Numbers*, fifth edition, John Wiley Sons, Inc.
- [11] G. Pólya, G. Szegő, *Problems and Theorems in Analysis*, Vol. I, Springer-Verlag (New York), 1972.

-
- [12] G. Pólya, G. Szegő, *Problems and Theorems in Analysis*, Vol. II, Springer-Verlag (New York), 1976.
 - [13] K. H. Rosen, *Elementary Number Theory and Its Applications*, Addison-Wesley (Reading), 1984.
 - [14] W. Sierpinski, *Elementary Theory of Numbers*, Polski Academic Nauk, Warsaw, 1964.
 - [15] W. Sierpinski, *250 Problems in Elementary Number Theory*, American Elsevier Publishing Company, Inc., New York, Warsaw, 1970.
 - [16] J. P. Serre, *A Course In Arithmetic*, Springer-Verlag (New York), 1973.

Other Books from XYZ Press

1. Andreescu, T., Elliott, S., *114 Exponent and Logarithm Problems from the AwesomeMath Summer Program*, 2017.
2. Andreescu, T., Crişan, V., *Mathematical Induction. A powerful and elegant method of proof*, 2017.
3. Andreescu, A., Andreescu, T., Mushkarov, O., *113 Geometric Inequalities from the AwesomeMath Summer Program*, 2017.
4. Bosch, R., *Cuban Mathematical Olympiads (2001-2016)*, 2017.
5. Matei, V., Reiland, E., *112 Combinatorial Problems from the AwesomeMath Summer Program*, 2016.
6. Andreescu, T., Mortici, C., Tetiva, M., *Pristine Landscapes in Elementary Mathematics*, 2016.
7. Andreescu, A., Vale, V., *111 Problems in Algebra and Number Theory*, 2016.
8. Andreescu, T., *Mathematical Reflections - two special years*, 2016.
9. Andreescu, T., Pohoata, C., Korsky, S., *Lemmas in Olympiad Geometry*, 2016.
10. Mihăilescu, C., *The Geometry of Remarkable Elements. Points, lines, and circles*, 2016.
11. Andreescu, T., Pohoata, C., *110 Geometry Problems for the International Mathematical Olympiad*, 2015.
12. Andreescu, T., Boreico, I., Mushkarov, O., Nikolov, N., *Topics in Functional Equations*, 2nd edition, 2015.
13. Andreescu, T., Ganesh, A., *109 Inequalities from the AwesomeMath Summer Program*, 2015.

14. Andreescu, T., Pohoata, C., *Mathematical Reflections - two great years*, 2014.
15. Andreescu, T., Ganesh, A., *108 Algebra Problems from the AwesomeMath Year-Round Program*, 2014.
16. Andreescu, T., Kisačanin, B., *Math Leads for Mathletes – a rich resource for young math enthusiasts, parents, teachers, and mentors*, Book 1, 2014.
17. Becheanu, M., Enescu, B., *Balkan Mathematical Olympiads – the first 30 years*, 2014.
18. Andreescu, T., *Mathematical Reflections – two more years*, 2013.
19. Andreescu, T., Rolinek, M., Tkadlec, J., *107 Geometry Problems from the AwesomeMath Year-Round Program*, 2013.
20. Andreescu, T., Rolinek, M., Tkadlec, J., *106 Geometry Problems from the AwesomeMath Summer Program*, 2013.
21. Andreescu, T., *105 Algebra Problems from the AwesomeMath Summer Program*, 2013.
22. Andreescu, T., Kane, J., *Purple Comet Math Meet! - the first ten years*, 2013.
23. Andreescu, T., *Mathematical Reflections - the next two years*, 2012.
24. Andreescu, T., Dospinescu, G., *Straight from the Book*, 2012.
25. Andreescu, T., *Mathematical Reflections – the first two years*, 2011.
26. Andreescu, T., Dospinescu, G., *Problems from the Book*, 2nd edition, 2010.